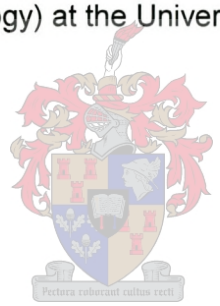
The molecular evolution of the spiral-horned antelope (Mammalia: Tragelaphini)

Sandi Willows-Munro

Thesis presented in partial fulfillment of the requirements for the degree of Master of Science (Zoology) at the University of Stellenbosch



Supervisors: Dr C. A. Matthee and Professor T. J. Robinson

December 2003

Declaration

I, the undersigned, hereby declare that the work contained in this thesis is my own original work and that I have not previously in its entirety or in part submitted it at any university for a degree.

Signature: **Date**:

Abstract

The evolutionary history of the African tribe Tragelaphini (spiral-horn antelope) is controversial. Past phylogenetic relationships among species were based on morphology or limited fossil evidence and are in conflict with mitochondrial DNA sequencing studies that have been conducted more recently. Although the group is distinguished from other African ungulates by the presence of spirally-twisted horns, the nine recognized extant species differ considerably in morphology, feeding habits and their habitat preference. The present study aims to resolve the phylogenetic uncertainties of the Tragelaphini using nuclear DNA sequence data derived from four independent DNA loci (MGF, PRKCI, SPTBN and THY). These data were combined with all previously published DNA sequences to produce a molecular supermatrix comprising approximately 6000 characters. Both parsimony and model based phylogenetic analyses of the nuclear DNA support the associations resulting from the analysis of mitochondrial genes. These findings suggest that the morphological characters previously used to delimit species within the group are subject to convergent evolution. The molecular phylogeny presented herein suggests that early members of Tragelaphini diverged from the other bovids during the mid-Miocene approximately 15.7 million years before present (MYBP). The common nyala (*Tragelaphus angasi*) and lesser kudu (*Tragelaphus imberbis*) representing the most basal species, separated from the other tragelaphids approximately 7.1 MYBP. This was subsequently followed by the radiation of those species adapted to a more tropical environment and they include the mountain nyala (*Tragelaphus buxtoni*), bongo (*Tragelaphus euryceros*), sitatunga (*Tragelaphus spekei*) and bushbuck (*Tragelaphus scriptus*), and the arid adapted clade comprising the giant eland (*Taurotragus derbianus*), common eland (*Taurotragus oryx*) and greater kudu (*Tragelaphus strepsiceros*). It is thought that this split occurred at the Miocene-Pliocene boundary approximately 5.4 MYBP. The timing of evolutionary events within the tribe suggests climatic oscillations and subsequent biotic shifts as the major driving forces underpinning speciation in the tribe Tragalaphini.

Opsomming

Die evolusionêre geskiedenis van die ras Tragelaphini (spiraalhoringwildsbokke) is kontroversieël. Vorige filogenetiese verwantskappe tussen die spesies is gebaseer op morfologie of beperkte fossiel bewyse. Meer onlangse studies, gebaseer op mitochondriale DNS nukleotieddata, is in teenstryding met baie van die evolusionêre hypotese afkomstig van morfologiese studies. Alhoewel die groep van die ander hoefdiere uitgeken kan word deur die aanwesigheid van spiraalvormige horings, verskil die nege hedendaagse spesies grootliks ten opsigte van morfologie, voedingswyse en habitat. Die hoof doelwit van hierdie studie was om die filogenetise verwantskappe tussen die Tragelaphini spesies te ontleed deur gebruik te maak van nukluêre DNS nukleotieddata afkomstig van vier onafhanklike DNS merkers (MGF, PRKCI, SPTBN en THY). Die data verkry is saamgevoeg by vorige gepubliseerde DNS nukleotidedata om 'n "supermatris" van sowat 6000 karakters te produseer. Parsimonie en modelgebaseerde filogenetise analise van die nukluêre DNS nukleotieddata het ooreengestem met die resultate van vorige mitochondriale studies. Hierdie bevindings dui daarop dat die morfologiese karakters wat voorheen gebruik is om die evolusionêre verwantskappe tussen die Tragelaphini spesies te ontleed onderhewig is aan konvergente evolusie. Die molekulêre filogenie wat hierin beskryf word stel voor dat die ras Tragelaphini gedurende die mid-Miocene, omtrent 15.7 miljoen jaar (MJ) gelede van die ander lede van die subfamilie Bovinae geskei het. *Tragelaphus angasi* en *Tragelaphus imberbis*, die mees basale spesies in die filogenie, het omtrent 7.1 MJ gelede van die ander lede van die Tragelaphini geskei. Hierdie skeiding is gevolg deur 'n split tussen die spesies aangepas vir 'n meer tropiese habitat (*Tragelaphus buxtoni*, *Tragelaphus euryceros, Tragelaphus spekei* en *Tragelaphus scriptus*) en die spesies aangepas vir 'n droë habitat (*Taurotragus derbianus, Taurotragus oryx* en *Tragelaphus strepsiceros*) Hierdie finale skeiding het gedurende die Miocene-Pliocene oorgang plaasgevind. Die tydsberekening van die evolusionêre gebeurtenisse wat binne die Tragelaphini ras plaasgevind het, gekoppel aan paleoklimaatdata, dui aan dat veranderinge in

klimaat en die geassosieerde habitatveranderinge verantwoordelik was vir die spesiasie patroon wat ons vandag in die Tragelaphini ras waarneem.

Acknowledgements

I am greatly indebted to everyone who contributed in different ways to the completion of this project. I would, in particular, like to thank my supervisors Conrad Matthee and Terry Robinson for providing direction and encouragement. My sincere appreciation to the staff and students of the evolutionary genomics laboratory at the University of Stellenbosch Zoology Department for both analytical and laboratory advice. I would also like to thank Carel van Heerden for sequencing assistance. This work would not have been possible without the financial assistance provided by the National Research Foundation and the University of Stellenbosch. Material was kindly provided by E. Harley (University of Cape Town), M. Harman (Powell Cotton Museum), J. Sakwa (University of Pretoria) and the Pretoria Zoological Gardens.

A special thanks must go to my family Mom, Dad, Megan and Gay for their continued loving support as well as my friends and colleagues Gavin Gouws, Amanda Pardini, Victor Rambau, Krystal Tolley and Peter Teske for all their invaluable advice and inspiration.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

Higher level taxonomy

The hollow-horned ruminantes, commonly referred to as bovids, have traditionally been placed within the order Artiodactyla (Schwabland 1993), renamed Cetartiodactyla following the inclusion of Cetacea as suggested by recent molecular studies (Graur and Higgins 1994, Gatesy *et al.* 1996, Montgelard *et al.* 1997, Gatesy *et al.* 1999, Matthee *et al.* 2001). The earliest fossil evidence (Gray 1821) does not distinguish between an Eurasian or African origin for the family Bovidae (Vrba 1985a, Hassanin and Douzery 1999a). A recent molecular study, however, indicates that the early members of the family probably arose in Eurasia during the early Miocene, approximately 20 million years before present (MYBP) from where they subsequently colonized the open niches in Africa (Matthee and Davis 2001).

At present the family is thought to comprise 49 extant genera and approximately 140 species (Vrba 1985a, Kingdon 1989, Nowak 1999). The majority of these occur on the African continent, although certain lineages are endemic to North America (e.g. *Bison bison*) and Eurasia (e.g. *Saiga tatarica*, Wilson and Reeder 1993). Although equivocal on morphological grounds (Janis and Scott 1987), the monophyletic status of Bovidae is well supported by molecular studies (Allard *et al.* 1992, Modi *et al.* 1996, Hassanin and Douzery 1999a, Matthee *et al.* 2001, Matthee and Davis 2001).

Traditionally the Bovidae is thought to comprise six subfamilies (Gentry 1992): Bovinae (cattle, buffalo and spiral-horned antelope), Cephalophinae (duikers), Hippotraginae (sable and roan antelope), Alcelaphinae (wildebeest and impala), Antilopinae (gazelle and dik-dik group) and Caprinae (sheep and goats). In contrast, investigations using DNA markers (Gatesy *et al.* 1997, Hassanin and Douzery 1999a, Hassanin and Douzery 1999b and Matthee and Davis 2001), chromosomes (Buckland and Evans 1978, Gallagher and Womack 1992, Robinson *et al.* 1998), allozymes (Georgiadis *et al.* 1990), serum immunology (Lowenstein 1986), physiological characters, and

morphological studies based on gland structure and horn type (Kingdon 1997) almost invariably support the delimitation of only two subfamilies. These are the Bovinae including the tribes Bovini, Boselaphini and Tragelaphini, and Antilopinae comprising the Alcelaphini, Hippotragini, Cephalophini, Reduncini, Antilopini, Aepycerotini, Oreotragini and Neotragini (Kingdon 1997, Bronner *et al.* 2003).

Taxonomy of the tribe Tragelaphini

The tribe Tragelaphini form a monophyletic lineage within the subfamily Bovinae (Essop *et al.* 1997, Matthee and Robinson 1999a, Hassanin and Douzery 1999a, Matthee *et al.* 2001, Matthee and Davis 2001). Members of the tribe may be morphologically distinguished from other representatives of the subfamily by the absence of internal sinuses in the frontals (Gentry 1978), and the presence of unringed, spirally-twisted horns. Several dental characteristics have been reported to unite the tragelaphid taxa, specifically the presence of upper molars without prominent lateral ribs between the styles. Importantly, however, this character is convergent when considering the closely related tribe Boselaphini which shares a similar dental characteristic (Gentry 1978).

Closely related sister tribes to the Tragelaphini include Boselaphini, containing the primitive Indian nilgai (*Boselaphus tragocamelus*) and the four-horned antelope (*Tetracerus quadricornis*), and the Bovini divided into buffalo Bovini (*Bubalus* and *Syncerus*) and cattle Bovini (*Bos* and *Bison;* Pitra *et al.* 1997). Both fossil and morphological evidence suggest that the Tragelaphini are a sister group of Boselaphini (Gentry 1990). Recent studies based on the combined analysis of multiple DNA fragments (Matthee *et al.* 2001; Hassanin and Douzery 2003), however, suggests a closer evolutionary relationship between the tribes Bovini and Tragelaphini, while the Boselaphini appears to be most basal in the Bovinae phylogeny. Given the limited taxon representation used by both studies, however, these findings must only be regarded as tentative (see Gatesy *et al.* 1997, Hassanin and Douzery 1999a).

The spiral-horned antelope are restricted to Africa and are divided into nine extant species (Figure I): greater kudu (*T. strepsiceros*), lesser kudu (*T. imberbis*), common nyala (*T. angasi*), mountain nyala (*T. buxtoni*), sitatunga (*T. spekei*), bushbuck (*T. scriptus*), common eland (*T. oryx*), derby or giant eland (*T. derbianus*) and bongo (*T. euryceros*) (Ansell 1971). The horns of the different species are distinct and permit easy species identification (Nowak 1999). The presence of horns in both sexes of *T. euryceros*, *T. oryx* and *T. derbianus* have been put forward as evidence that these taxa are a derived monophyletic entity (Gentry 1990). This finding is not, however, supported by studies based on general body form and pelt colouration which unite the eland (*T. derbianus* and *T. oryx*) and kudu (*T. strepsiceros* and *T. imberbis*, Walker 1964) lineages. Haltenorth (1963) warned against the usage of these latter characters for phylogeny reconstruction as they are extremely variable even between subspecies. Cranial similarities have also in the past been used to unite *T. spekei* and *T. angasi* (Roberts 1952, Kingdon 1989), and *T. imberbis* and *T. strepsiceros* (Kingdon 1989, Alden *et al.* 1995). The unreliable nature of these characteristics is also evident when considering *T. buxtoni*. The latter was first described as intermediate between the *T. strepsiceros* and *T. angasi* and originally named *Strepsiceros buxtoni* by Lydekker (1910). Only a year later Lydekker (1911) reconsidered this association favoring an affinity between this species and *T. angasi* and renamed the species *T. buxtoni*. The inconclusiveness of many of the classic morphological studies has lead many researchers to regard bovids as one of the most difficult of all mammalian families to classify (Simpson 1945) which, given the focus of this study, lead Ansell (1971) to state that the "classification of this group [tragelaphids] has caused considerable divergence of opinion".

Three tragelaphid genera have historically been recognized. *Tragelaphus* containing *T. strepsiceros*, *T. imberbis*, *T. angasi*, *T. buxtoni*, *T. spekei* and *T. scriptus*; *Taurotragus* containing *T. oryx* and *T. derbianus*; and *Boocercus* for *T. euryceros*. It has been suggested that the genus *Boocercus* be considered a subgenus of *Tragelaphus* (Ansell 1971, Wilson and Reeder 1993). The close relationship between these two lineages suggested by mitochondrial sequence evidence (e.g. Hassanin and Douzery 1999a, Matthee and

Cont…

Figure I: Distributions of a) *T. imberbis*, b) *T. angasi*, c) *T. buxtoni,* d) *T. scriptus*, e) *T. spekei*, f) *T. euryceros,* g) *T. derbianus*, h) *T. oryx*, i) *T. strepsiceros* redrawn from Schwabland 1993.

Cont…

Figure I: Distributions of a) *T. imberbis*, b) *T. angasi*, c) *T. buxtoni,* d) *T. scriptus*, e) *T. spekei*, f) *T. euryceros,* g) *T. derbianus*, h) *T. oryx*, i) *T. strepsiceros* redrawn from Schwabland 1993.

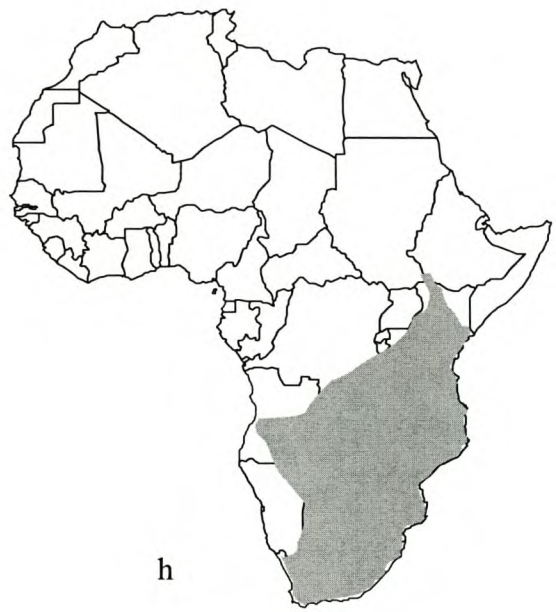


Figure I: Distributions of a) *T. imberbis*, b) *T. angasi*, c) *T. buxtoni,* d) *T. scriptus*, e) *T. spekei*, f) *T. euryceros,* g) *T. derbianus*, h) *T. oryx*, i) *T. strepsiceros* redrawn from Schwabland 1993.

Robinson 1999a) is further underscored by the ability of *T. euryceros* and *T. spekei* to hybridize even though these two species differ extensively in morphology (Tijskens 1968). The taxonomic status of the genus *Taurotragus* has similarly been questioned. Ellerman *et al.* (1953) placed *T. derbianus* and *T. oryx* in a separate genus given the marked behavioral and morphological differences which include horns in an open spiral and the absence of a dewlap (Nowak 1999). In contrast to this Haltenorth (1963) argued for the placement of all nine extant species of Tragelaphini in a single genus, *Tragelaphus*, with *Taurotragus* as a subgenus. At present the recognition of a single genus enjoys support from numerous studies based on fossil evidence (Vrba 1987), mitochondrial DNA sequences (Essop *et al.* 1997, Gatesy *et al.* 1997, Hassanin and Douzery 1999a, Matthee and Robinson 1999a), nuclear DNA sequences (Hassaini and Douzery 1999b) and allozymes (Geogiadis *et al.* 1990).

Choice of genetic markers

A number of studies based on mitochondrial DNA regions (Gatesy *et al.* 1997, Hassanin and Douzery 1999a, Hassanin and Douzery 1999b, Matthee and Robinson 1999a) have included tragelaphid species. Unfortunately all studies had incomplete taxonomic representation of the tragelaphids, so the conclusions pertaining to Tragelaphini were limited. The most taxonomically conclusive study, an analysis of cytochrome *b* sequences by Matthee and Robinson (1999a) supported the close evolutionary relationship between *T. euryceros* and *T. spekei*, invalidating the generic status of *Boocercus*. This view enjoys support from other studies using the same marker (Hassanin and Douzery 1999a, Hassanin and Douzery 1999b). Similarly, the phylogenetic relationship suggested by cranial similarities of *T. imberbis* and *T. strepsiceros* (Kingdon 1989, Alden *et al.* 1995), and *T. angasi* and *T. spekei* (Roberts 1952, Kingdon 1989) was not supported by the analysis of mitochondrial DNA sequence data.

Matthee and Robinson (1999a) suggested that the contrasting phylogenetic associations evidenced by the mitochondrial DNA and the morphological data may be due to convergent morphological evolution, and that speciation in the

group was probably the result of vicariance. This was based on the close relationship between *T. euryceros*, *T. spekei* and *T. scriptus* observed in mitochondrial DNA and it was argued that these seemingly derived species all occur in dense forest habitat (Nowak 1999). The other basal species *T. imberbis*, *T. angasi*, *T. strepsiceros*, *T. oryx* and *T. derbianus*, are all arid adapted suggesting that tragelaphids may have evolved in the drier bushlands and savannas of Africa and only subsequently moved into moist woodlands and forest habitats (Hassanin and Douzery 1999a, Hassanin and Douzery 1999b, Matthee and Robinson 1999a). This proposed direction of evolutionary radiation is in contrast to that proposed by Vrba (1987) and Kingdon (1989) who placed *T. spekei* basal in the Tragelaphini radiation. A similar hypothesis was suggested by an allozyme study (Geogiadis *et al.* 1990), which placed *T. scriptus* basal to the kudu and eland in all trees, a conclusion supported by restriction site mapping of mitochondrial DNA (Essop *et al.* 1997). If this holds, the basal placement of the *T. spekei*, *T. euryceros* and *T. scriptus* clade suggests that the tribe Tragelaphini evolved in moist forests and only later became adapted to drier grasslands. Such a hypothesis is in close agreement with the extreme climatic changes suggested to have occurred during the Miocene resulting in the expansion of savanna woodlands and grasslands at the expense of forest habitats (Zachos *et al.* 2001). These changes created conditions favoring the evolution of grazing and thus the more arid adapted species.

In addition to mitochondrial DNA sequences, data for a variety of nuclear loci are accumulating for most animal groups. Substantially larger than the mitochondrial genome (approximately 166,000 times larger), the mammalian nuclear genome can provide a suite of independently segregating markers (Matthee *et al.* 2001). A number of recent studies (Gatesy and Arctander 2000, Madsen *et al.* 2001, Matthee *et al.* 2001, Murphy *et al.* 2001, Springer *et al.* 2001) highlight the usefulness of nuclear DNA in resolving systematic uncertainties. Although the majority of these studies have been based at higher taxonomic levels, nuclear DNA markers were considered to potentially have two benefits for the present study. First, variation in the evolutionary rate of different nuclear fragments may allow resolution of both the deepest

and most recent nodes of the evolutionary tree (Harrison 1991, Hillis and Dixon 1991, Hillis 1999). Second, highly supported nodes retrieved from the analysis of both the independent and combined nuclear data could be used as supporting evidence for the validity of either the mitochondrial DNA or morphological phylogenies.

Gene trees versus species trees

The development of the polymerase chain reaction (PCR, Mullis and Faloona 1987) revolutionized the field of molecular biology and had an immense impact on phylogeny reconstruction. Given the availability of universal mammalian primers (Kocher *et al.* 1989), the high mutation rate in animals (Brown *et al.* 1982) and ease of alignment among lineages (Irwin *et al.* 1991), mitochondrial DNA markers have been among the most exploited genomic regions. Although mitochondrial DNA has performed well in recovering phylogenies at nearly all taxonomic levels (Manceau *et al.* 1999, Matthee and Robinson 1999b), the lack of congruence between the mitochondrial and morphological phylogenies obtained for the tribe (Hassanin and Douzery 1999a, Matthee and Robinson 1999a) may be attributed to the rapid evolutionary radiation of the tragelaphids in Africa. Such rapid cladogenesis would have allowed little time for synapomorphies to accumulate thereby reducing the robustness of clades detected in phylogenetic analyses (Gatesy *et al.* 1992). Similar scenarios have been proposed for many mammalian groups for example groups within the Cetartiodactyla (Allard *et al.* 1992, Matthee and Robinson 1999a, Rebholz and Harley 1999), Rodentia (Lara *et al.* 1996, Matthee and Robinson 1997), Carnivora (Flynn and Nedbal 1998) and Logomorpha (Halanych *et al.* 1999). Another confounding effect influencing gene trees could be independent lineage sorting which can lead to increased homoplasy where polymorphic characters have become fixed in only a portion of the descendents of a polymorphic ancestor (Hillis 1999). Additionally, the faster rate of nucleotide substitutions in mitochondrial DNA loci may have resulted in multiple changes and subsequently to saturation leading to lack of resolution in these trees (e.g. Allard *et al.* 1992, Gatesy *et al.* 1997).

Given the problems effecting phylogenies derived from a single genetic marker, it is necessary to make a distinction between species trees and gene trees. A species tree represents the evolutionary history of a group, while a gene tree reflects the diversification of a group of sequences within the context of a species tree and as such is sensitive to the effects of hybridization, lineage sorting and gene duplication events (Moritz and Hillis 1996, Lyons-Weiler and Milinkovitch 1997, Nei and Kumar 2000). The collection of data from multiple and preferably independent markers may be used to circumvent the discrepancies between gene and species trees (Pamilo and Nei 1988, Atchley and Fitch 1991, Slade *et al.* 1994). Although the use of independent markers may not always detect gene tree and species tree conflicts, especially when comparing DNA regions that have undergone identical lineage sorting (Lyons-Weiler and Milinkovitch 1997), congruence among phylogenies derived from different data classes is considered the most reliable indication in support of a species tree.

The combined approach to phylogenetic estimation

Recent studies based on the combined analysis of nuclear and mitochondrial sequences (Gatesy and Arctander 2000, Remigio and Hebert 2000, Murphy *et al.* 2001, Birks and Edwards 2002, O'Grady and Kidwell 2002) have suggested that the rapid cladogenesis evidenced of many mammalian taxa may be resolved using a combination of data sets from multiple sources. Supertree and supermatrix approaches both allow for the simultaneous examination of phylogenetic information from multiple sources (Gatesy *et al.* 2002). Supertrees are constructed utilizing topologies as phylogenetic evidence with source trees recorded as a series of binary characters describing each node (Jones *et al.* 2002). Although useful in incorporating information from past studies where the raw character data is unavailable, the methodology used in supertree construction can easily be confused, and should thus be used with the necessary caution (Gatesy *et al.* 2002). Unlike supertrees, the supermatrix approach involves the merging of all raw character data into a single data matrix (Gatesy *et al.* 2002). The supermatrix approach allows for maximum interaction of character information and has the

potential to increase resolution, nodal support and accuracy (Remigio and Hebert 2000, Gatesy *et al.* 2002).

Aims and objectives of the study

Evolutionary processes are often complex and require an extensive understanding of the phylogenetic relationships within a group. To this end, an accurate phylogeny for the Tragelaphini is essential for interpreting the pattern of lineage turnover within the group. The various species differ in feeding (Gagnon and Chew 2000) and habitat preferences (Nowak 1999) and these, together with the extent and timing of episodes of speciation, should contribute to the broader understanding of environmental and climatic changes in Africa.

The aims of this study were threefold:

1. To investigate the ability of conservative nuclear DNA intron sequences to recover a robust phylogeny for a closely related group of bovids characterized by a recent radiation.
2. To refute or corroborate the taxonomic discrepancies of the past, in particular the relationship between *T. spekei*, *T. scriptus* and *T. euryceros* observed in the mitochondrial DNA phylogenies. Similarly, an attempt was made to investigate the relationships between morphologically similar taxa (*T. imberbis* and *T. strepsiceros*; *T. angasi* and *T. spekei*).
3. To date the radiation events within the Tragelaphini, and correlate these data with paleoclimatic change in Africa.

Chapter 2: Materials and methods

Genomic DNA extraction and nucleotide sequencing

Zwickl and Hillis (2002) demonstrated that increased taxonomic sampling can improve phylogenetic resolution. Where possible, the present study is the first to include all nine recognized tragelaphid species and to use multiple representations thereof (Table I). Given the scarcity of some species and the difficulties encountered in obtaining fresh samples, in some cases dried museum skins were used for DNA extraction. Total genomic DNA was extracted using the standard phenol/chloroform/iso-amyl alcohol DNA protocol (Sambrook *et al.* 1989). Each extraction cycle was checked for contamination by including a negative extraction (i.e. all reagents and steps without DNA, Sambrook *et al.* 1989).

Standard polymerase chain reaction (PCR) procedures were followed. The PCR volume of 50μl contained 5μl (10x) buffer, 5μl $MgCl_2$ (25mM), 2μl of each primer (50μM), 10μl dNTPs (1mM), 24μl ultra-pure water and one unit super-therm DNA polymerase (Southern Cross Biotechnology). PCR amplification was performed under the following conditions: one cycle of 3min denaturing; 25 cycles, with temperature 95° for 30s; 50 ° to 58 °C for 30s; 72 °C for one minute and one last cycle at 72 °C for 5min (for further details see Matthee *et al.* 2001). Standard precautions including negative controls (template-free PCR reactions) were used to test for contamination of the PCR process, and to ensure the fidelity of PCR reactions. Both strands were sequenced to improve the accuracy of base identification. PCR products were separated in 0.8% agarose gels and purified using a commercially available extraction kit (Nucleospin Extract- Macherey-Nagel). Cycle sequencing was performed using Big Dye chemistry (Version 3, Applied Biosystems). The resulting products were analyzed on a 3100 ABI automated sequencer.

In an attempt to investigate the utility of nuclear markers for the tragelaphid evolution, four nuclear DNA regions were used. These include thyrotropin (THY), protein-kinase C iota (PRKCI), B-spectrin nonerythrocytic (SPTBN)

Table I: Tragelaphid species used in the study. Vernacular names, collection localities and source material used for DNA extractions are provided.

Species name	Vernacular name	Collection locality	Material	Material provided
Taurotragus derbianus	Giant eland	Central Africa	DNA	Pretoria Zoological Gardens
Taurotragus oryx	Common eland	East Africa	DNA	Pretoria Zoological Gardens
Taurotragus oryx	Common eland	Tanzania	DMSO preserved tissue	Unknown
Tragelaphus strepsiceros	Greater kudu	Unknown	DNA	Pretoria Zoological Gardens
Tragelaphus strepsiceros	Greater kudu	Namibia	Dried tissue	J. Sakwa, University of Pretoria
Tragelaphus imberbis	Lesser Kudu	East Africa	DNA	Pretoria Zoological Gardens
Tragelaphus angasi	Common nyala	Unknown	DNA	Pretoria Zoological Gardens
Tragelaphus angasi	Common nyala	South Africa	DMSO preserved tissue	Unknown
Tragelaphus buxtoni	Mountain nyala	Ethiopia	Dried tissue	Powell-Cotton Museum
Tragelaphus buxtoni	Mountain nyala	Ethiopia	Dried tissue	Powell-Cotton Museum
Tragelaphus euryceros	Bongo	Unknown	DNA	Pretoria Zoological Gardens
Tragelaphus euryceros	Bongo	Congo	DMSO preserved tissue	Unknown
Tragelaphus spekei	Sitatunga	Unknown	Tissue culture	E. Harley, UCT Medical School
Tragelaphus scriptus	Bushbuck	South Africa	Tissue culture	E. Harley, UCT Medical School
Tragelaphus scriptus	Bushbuck	Southern Africa	Dried tissue	Unknown

and stem cell factor (MGF). All primers used in the present investigation have been successfully used in previous studies on bovids including selected tragelaphid species (Matthee *et al.* 2001, Matthee and Davis 2001). Primers are located within conserved coding regions and were selected from a set of markers previously used to detect target inserts in bovine and ovine bacterial artificial chromosome (BAC) libraries (Matthee *et al.* 2001). Most nuclear sequence data were obtained from faster-evolving intron regions. An additional four mitochondrial DNA and two nuclear DNA regions as well as sequence data from outgroup taxa were obtained from public databases (Table II). These included 12S rRNA (Gatesy *et al.* 1992), 16S rRNA (Gatesy *et al.* 1992), cytochrome *b* (cyt *b*, Hassanin and Douzery 1999a, Matthee and Robinson 1999a), cytochrome oxidase subunit III (cox III, Rebholz and Harley unpublished) and nuclear regions aromatase cytochrome p-450 (cyp 19, Hassanin and Douzery 1999b) and lactoferrin encoding region (lactoferrin, Hassanin and Douzery 1999b). In an attempt to reduce missing data and with the availability of universal primers in the laboratory (see Matthee and Davis 2001 for primer details) we extended the Genbank data by sequencing all additional duplicate taxa included in our study. Unfortunately primers were not available in the laboratory to complete the cox III, cyp 19 and lactoferrin DNA data sets which were obtained from Genbank and taxa included in our study without these sequences were added to the supermatrix but with characters scored as missing.

Discussion of analytical methods used

i) Outgroup selection

Although it has been suggested that phylogenetic analysis can be performed, at least initially, with a single outgroup (Nixon and Carpenter 1993), the inclusion of more than one outgroup may reduce the introduction of errors in the polorization of characters and thereby limit the retrieval of erroneous evolutionary relationships (Milinkovitch *et al.* 1996, Milinkovitch and Lyons-Weiler 1998). Additionally the inclusion of multiple outgroups of successive relatedness to the ingroup reduces the effects of long-branch attraction in parsimony analysis (Smith 1994) and can be used to test the assumption of

Table II: Taxonomic sampling used in the study. Accession numbers are provided where sequence data was used from previous studies. * indicates specimens sequenced in the present study, the ■ indicates missing data.

	DNA Region									
	Mitochondrial DNA						Nuclear DNA			
Species name	12S	16S	Cox III	Cyt *b*	Cyp 19	Lactoferrin	MGF	PRKCI	SPTBN	THY
Tuarotragus derbianus	*	*	■	AF022062	■	■	*	*	*	*
Taurotragus oryx	AF091710	*	AF030275	AF036278	AF091675	AF091646	*	*	*	*
Taurotragus oryx	*	TOU87060	■	AF022057	■	■	*	*	*	*
Tragelaphus strepsiceros	*	*	AF030278	AF022063	AFO91676	AF091648	*	*	*	*
Trsgelaphus strepsiceros	AF091696	*	■	AF036280	■	■	*	*	*	*
Tragelaphus imberbis	*	TAGMTTGRG	AF030274	AF036279	AF091677	AF091649	*	*	*	*
Tragelaphus angasi	AF091698	*	AF030271	AF091633	AF091674	AF091650	*	*	*	*
Tragelaphus angasi	*	*	■	AF022066	■	■	*	*	*	*
Tragelaphus buxtoni	*	*	AF030272	*	■	■	*	*	*	*
Tragelaphus buxtoni	*	*	■	*	■	■	*	*	*	*
Tragelaphus euryceros	AF091691	*	AF030273	AF036276	AF091678	AF091644	*	*	*	*
Tragelaphus euryceros	*	*	■	AF022065	■	■	*	*	*	*
Tragelaphus spekei	AF091692	*	AF030277	AJ222680	AF091663	AF091645	*	*	*	*
Tragelaphus scriptus	*	*	AF030276	AF036277	AF091679	AF091647	*	*	*	*
Tragelaphus scriptus	AF091693	*	■	AF022067	■	■	*	*	*	*
Boselaphus tragocamels	BOSMTTGRG	BTU87013	■	AJ222679	AF091672	AF091642	AF165724	AF165725	AF165726	AF165729
Bos taurus	BOVMT	BOVMT	BOVMT	BOVMT	BTAROCG	AF281088	AF165716	AY029293	AF165718	BOVTSHB
Syncerus caffer	AF091688	SCU87061	AF030279	AF036275	AF091670	AF091640	*	AF210175	AF210197	AF210219
Aepyceros melampus	AEYMTTGRG	AEYMTTGRG	AF030456	AEYMTTGRG	AFO91680	AF091659	AF165780	AF165781	AF165782	AF165785

ingroup monophyly (Baverstock and Moritz 1996). For this reason four outgroups of successive relatedness to the ingroup were used in the phylogenetic analysis of the tribe Tragelaphini. The Indian nilgai (*Boselaphus tragocamelus*), cattle (*Bos taurus*) and African buffalo (*Syncerus caffer*) were used as representatives of the three currently recognized sister lineages (Boselaphini, cattle Bovini and buffalo Bovini). In addition the impala (*Aepyceros melampus*), a member of Antilopinae (following Bronner *et al.* 2003), was used to represent a distantly related lineage (Matthee and Davis 2001).

ii) Alignment

All sequences were aligned using the default settings of CLUSTAL X (Thompson *et al.* 1997) and were optimized manually to ensure homology (Appendix). Due to conservation of motifs and sequence length all sequences were easily aligned and ambiguously aligned sites were excluded (Appendix: 4045-4053). Any heterogeneous changes were coded using the IUB code. Exon and intron boundaries were defined from previously published sequence data for each DNA region (Matthee and Davis 2001, Matthee *et al.* 2001). To further ensure the fidelity of the DNA sequences, exon regions were translated into amino acids and screened for functionality. No stop codons or insertions were present in the exonic regions.

iii) Data partitioning

All characters and taxa were merged into a single data matrix or "supermatrix" containing four regions of mitochondrial DNA origin and six regions of nuclear DNA origin. For our supermatrix, a number of partitioning scenarios could be justified and these were analyzed independently. Each DNA region amplified was considered separately, or partitioned according to the origin of marker (either mitochondrial or nuclear). This approach was followed because all the mitochondrial genes are linked and represent a single marker system, and independent examination of each of these DNA regions or partitions could establish what effect each has on the supermatrix. The corroboration from many independent DNA regions and partitions is accepted as providing strong

evidence for the reliability of the resulting phylogenetic trees (Miyamoto and Fitch 1995).

iv) Congruence of different data partitions

The inclusion of DNA regions which show conflict in a supermatrix has been suggested to potentially mislead phylogenetic estimation (Bull *et al.* 1993). To explore inconsistencies in phylogenetic signal among the mitochondrial and nuclear partitions, incongruence length differences (ILD) were calculated in PAUP* 4.0b10 (Swofford 1999). The ILD or partition homogeneity test was originally designed to test the effect of data combination on homoplasy levels, and assumes that as homoplasy increases, phylogenetic accuracy decreases (Yoder *et al.* 2001). Following Lee's (2001) suggestion that the inclusion of uninformative characters may cause the ILD test to overestimate the amount of incongruence between DNA regions, all uninformative characters were excluded prior to ILD analysis. It is important to realize that although the ILD test has in the past been widely accepted as an effective statistical method for deciding whether to combine independent data sets (Yoder *et al.* 2001), numerous studies have cautioned against the use of ILD. This is because homoplasy may have virtually no correlation with phylogenetic accuracy (Sanderson and Donoghue 1989, Yoder *et al.* 2001). Additionally, lack of congruence may be a result of missing taxa and the limited number of informative characters in some DNA regions (Gatesy and Arctander 2000).

To further explore the heterogeneity among the DNA regions as well as the designated partitions, a Bayesian approach similar to that previously applied by Buckley *et al.* (2002) was also taken. Buckley *et al.* 2002 estimated the 0.95 posterior interval (the set of topologies contained in 0.95 of the posterior distribution) for a number of DNA regions using MrBayes 2.01 (Huelsenbeck and Ronquist 2001) and the summarise option in BAMBE 2.03 (Simon and Larget 2000). A maximum likelihood topology for the combined data sets was estimated, assuming the same model for all partitions. Limited signal heterogeneity was assumed if the maximum likelihood topology fell within the 0.95 posterior interval for each DNA region. Although successfully applied to New Zealand *Circada* data (Buckley *et al.* 2002), for the tragelaphids the

paucity of variable characters in many independent DNA regions resulted in several unresolved nodes in our study meaning that the supermatrix maximum likelihood topology would almost never fall within the 0.95 Bayesian posterior interval for a number of DNA regions. The original method employed by Buckley *et al.* (2002) would thus incorrectly assume that these unresolved data sets contain signal conflict. To overcome this problem we simply compared topological congruence visually. Additionally, incomplete taxonomic sampling characterizing some DNA regions, in particular those obtained from previously published studies (i.e. cox III, cyp 19 and lactoferrin) required that we performed the maximum likelihood analyses on truncated data sets. Lack of conflict between 0.95 Bayesian and maximum likelihood trees would suggest that the observed topologies could have given rise to the sequence data. Conversely incongruence among trees would indicate that a particular DNA region or partition evolved along an alternative topology or, alternatively, the selected substitution model was miss-specified (Buckley *et al.* 2002).

v) Selecting the best-fit model of nucleotide substitution

Numerous studies have cautioned against the use of an inappropriate model of evolution in phylogenetic analysis (Sullivan and Swofford 1997, Cunningham *et al.* 1998, Kelsey *et al.* 1999) and in particular the negative effects on branch length and bootstrap support are well documented (Yang 1994, Buckley *et al.* 2001). The choice of model should be justified using statistical tests and a number of ways to do this have been suggested. The likelihood ratio test (LRT) has been extensive used in the phylogenetic literature but the use of Akaike information criterion (AIC, Akaike 1974) and Bayesian information criterion (BIC, Schwarz 1974) is also well entrenched (Hasegawa 1990, Tamura 1994, Morozov *et al.* 2000). MODELTEST 3.06 developed by Posada and Crandall (1998) utilizes both the AIC and LRT to select an appropriate model of evolution. The LRT compares the maximum likelihood value of a topology estimated under a simple model, with the optimal topology using a more complex model. The more parameter rich model is only accepted if the addition of parameters significantly improves the fit of the evolutionary model to the data. The AIC, which is based on an

estimate of the Kullback-Leibler (K-L) distance, gives an indication of the expected distance between any model and the evolutionary process which generated the data (Posada and Crandall 2001). While both the LRT and AIC compare competing models, Posada and Crandall (2001) found that the model selection process was, in some cases, influenced by the order in which models were tested, and the complexity of the initial model in LTR test. Despite this, Posada and Crandall (2001) suggest that the LRT generally performed better than the AIC under various conditions. Based on these findings, and the concordance of the LTR and AIC tests, the evolutionary model selected by the LRT for each DNA region or partition in the present study was used in all maximum likelihood and Bayesian analyses.

vi) Phylogenetic estimation

Parsimony and maximum likelihood phylogenies were derived for individual DNA regions, the mitochondrial partition, the nuclear partition and the supermatrix using PAUP* 4.0b10 (Swofford 1999). Parsimony and maximum likelihood analysis were performed using the heuristic search option with 100 replicates of random taxon addition and TBR branch swapping (PAUP* 4.0b10, Swofford 1999). The consistency index (CI, Klug and Farris 1969) and retention index (RI, Farris 1989) was used to access the relative amounts of homoplasy present in each DNA region and partition. Although both the CI and RI measure the amount of homoplasy in a data set, the RI is considered the more reliable indicator of saturation as it is less likely to vary with topology (Nei and Kimar 2000). The use of differential weighting systems has been extensively advocated in the literature (Goloboff 1993, Swofford *et al.* 1996, Broughton *et al.* 2000) as a way of improving resolution by down-weighting characters considered noisy or misleading (i.e. transitional changes). Similarly, coding gaps as a 5th character (i.e. weighting indels proportional to size) has been considered an effective way of incorporating phylogenetic information contained within indels (e.g. Gatesy and Arctander 2000). Analyses were performed on the mitochondrial partition, the nuclear partition and the supermatrix using differential weighting schemes and coding gaps as a 5th character. Character transformations were also equally weighted with gaps treated as missing data.

vii) Nodal stability

The reliability of a phylogeny can only be assessed by testing the relative stability of nodes (Sanderson 1995). In the present investigation, the stability of nodes was assessed using non-parametric bootstrap (Felsenstein 1985) and decay index values (Bremmer 1988, 1994). Bootstrap analysis involves assembling replicate data sets the same size as the original data matrix, but containing randomly sampled characters from the same data set (Felsenstein 1985). Nodal support for parsimony trees was assessed by 1000 bootstrap iterations (Felsenstein 1985). Computational limitations allowed only 100 replications for maximum likelihood analysis. Decay index values were estimated using Autodecay 4.0 (Eriksson 1998) which provides the number of additional steps or synapomorphic nucleotide changes required to collapse a specific node (i.e. the length difference between the most-parsimonious tree, which retrieves a node and the best tree lacking that node, Bremmer 1988, 1994). Support for nodes was also assessed using posterior probability values generated by the Bayesian analysis (see below).

viii) Distribution of support among different data partitions

Conventional tests for data incompatibility (i.e. the ILD and Bayesian approach described previously) measure the overall conflict among data sets and are therefore not sensitive enough to detect conflict between particular clades (Lee and Hugall 2003). The influence of different DNA regions on selected nodes was assessed in a number of ways. First, following Kluge's (1989) suggestion that the spread of synapomorphic characters among data sets is a good reflection of the support given by each DNA region to the supermatrix topology, unambiguous synapomorphies were derived for each node supported by the supermatrix analysis. This was done using the trace all characters option in MacClade 4.0 (Maddison and Maddison 1992). Second, nodal data set influence (NDI, e.g. Gatesy *et al.* 1999) for selected nodes were calculated for each of the 10 DNA regions. NDI measures the support given by each DNA region for selected nodes in the supermatrix analysis and it identifies potential areas of conflict between data sets (Lee and Hugall 2003). NDI represents the decay index value at a particular node for the supermatrix minus the decay index value at that node for the supermatrix

excluding a specified DNA region. Scores may be positive, negative, or zero, and reflect the effect that the removal of a specific DNA region has on the level of support at a selected node. Following Gatesy *et al.* (1999) a positive NDI value (in our analyses indicated in white, see later) indicates that removal of a particular DNA region reduced support at a selected node. In other words, this indicates that the particular DNA region contains phylogenetic signal in support of a node. Conversely, a DNA region responsible for a negative NDI at a node (in our analyses indicated in black) is assumed to have had a negative effect on the selected node by increasing support once that DNA region has been removed from the supermatrix. An NDI of zero (in our analyses indicated in grey) indicates that a DNA region had no influence on the support at a particular node.

As with decay index values proposed by Bremer (1988, 1994), the significance of NDI values is highly dependent on the data examined and may vary between phylogenies due to differences in the number of characters examined. This may also be applicable within phylogenies due to the differential distribution of homoplasy (Lee 2000). The application of a standard significance level is therefore inappropriate (Lee and Hugall 2003) thus making comparisons between data sets problematic (Bremer 1988, Sanderson and Donoghue 1996). Lee (2000) proposed a method of assessing nodal significance using a method similar to that used in decay analysis. The most-parsimonious tree containing a node of interest and the most-parsimonious tree without that node are compared. However, unlike decay analysis where the additional steps or nucleotide changes required to collapse a specific node are considered, AutoCladeS 0.2 (Ericksson 2001) uses Templeton's test (1983) in PAUP* 4.0b10 to statistically compare individual nodes on alternative trees. If the most-parsimonious tree is better supported than the constraint tree ($P \leq 0.1$) the node is considered significantly supported. Alternatively, if the constraint tree is favored ($P > 0.1$) the node is considered unsupported. The statistical quantification of support at nodes facilitates comparisons among and within data sets, allowing for the identification of hidden character support or conflict by the independent DNA regions. For example, DNA regions not capable of retrieving a specified node

when independently examined may lend support, or conflict, to a node within the context of the supermatrix (Gatesy *et al.* 1999). Truncated data sets (i.e. supermatrix lacking a particular DNA region) were analyzed using AutoCladeS 0.2 in combination with the Templeton's test in PAUP* 4.0b10 to statistically examine the effect the removal of a specified DNA region has on a particular node. Following Lee (2000) who considered the 0.05 confidence limit to be excessively conservative, we considered nodes with $P \leq 0.1$ significantly supported.

ix) Bayesian inference

A number of authors have noted the superiority of model-based approaches in phylogeny construction (Huelsenbeck and Hillis 1993, Kuhner and Felsenstein 1994, Huelsenbeck 1995). Lack of computational resources, however, makes conventional maximum likelihood analysis on large data sets difficult. Bayesian inference of phylogeny has recently been suggested as a practical alternative to conventional maximum likelihood based methods. Bayesian estimation of phylogeny rests on the posterior probability of the distribution of trees. This is estimated using the Markov Chain Monte Carlo (MCMC) method (Mau 1996, Rannala and Yang 1996, Mau and Newton 1997, Yang and Rannala 1997, Mau *et al.* 1999). The method uses the fraction of time that a Markov chain visits any particular tree as a valid approximation of the posterior probability of that tree. Unlike conventional bootstrap techniques, Bayesian estimation does not require the time-consuming re-optimization of the likelihood function over many replicates (Felsenstein 1985, Goldman *et al.* 2000).

Although computationally effective on large data sets, a number of practical problems related to the Bayesian method have to be addressed. A Markov chain is typically started from a random point which may be far removed from regions within the parameters space containing trees with the highest posterior probabilities. A chain often requires a period of time before finding trees representing a reasonable approximation of the posterior distribution. Additionally, there is always the chance that any single chain may never find the area representing the optimal tree space. To overcome this problem each

Bayesian analysis was run for at least $1X10^6$ generations with trees sampled every 100^{th} generation resulting in a total of 10 000 trees saved per simulation. To check that stability (i.e. lack of improvement in maximum likelihood scores) had been reached, each simulation was repeated three times and variation in the maximum likelihood scores were visualized graphically using Microsoft ® Excel (2002). Once stability was determined, the topologies that formed part of the initial climbing stage (referred to as the "burnin") were removed. This is normally obtained no later than 20 000 simulations. Additionally, the MCMC process is set so that four chains (three heated and one cold) are started from random trees and run simultaneously. To further increase the chance that the chains would converge on the area with the highest posterior probability values, in each generation two chains were picked randomly and their states changed so the cold chain, the source from which trees are sampled, was provided with proposals of new states from the hot chains.

The maximum likelihood parameters for initial Bayesian runs performed by MrBayes 3.0b3 (Huelsenbeck and Ronquist 2001) on the independent DNA regions, the mitochondrial partition, the nuclear partition and the supermatrix were set as follows, "lset nst = 6" (the general time reversible model, Rodríguez *et al.* 1990), "rates = invgamma" (site-specific rate variation such that some sites are invariant and the rates of others are drawn from the gamma distribution), "basefreq = estimate" (estimated proportion of base types from the data). Additionally, analysis were performed using the partition function available in MrBayes 3.0b3. This differentially applies evolutionary models to selected DNA regions allowing the evolutionary constraints of the different markers to be incorporated into an analysis. The posterior probability values used as a statistical means of measuring nodal confidence in Bayesian inference have been found, in many cases, to be higher than conventional non-parametric bootstrap values (also see Suzuki *et al.* 2002). This difference has often been interpreted as indicating that conventional bootstrapping methods are too conservative (Hillis and Bull 1993, Murphy *et al.* 2001). Recent studies, however, suggest that posterior probability values may in fact be inflated (Waddell *et al.* 2001, Suzuki *et al.* 2002) and consequently

confidence was given to nodes consistently supported by both methods (bootstrap ≥ 70% and posterior probability values ≥ 95%).

x) Tests of alternative molecular phylogenies

Conflicting associations derived from the mitochondrial partition, nuclear partition and supermatrix were explored using constraint analysis. To determine if an alternative topology differed significantly from a best estimate of phylogeny, constraint analysis was performed using the Shimodaira-Hasegawa test (S-H test, Shimodaira and Hasegawa 1999) in PAUP* 4.0b10. If the log likelihood value from the constraint topology was significantly greater than the log likelihood value from the best topology than the alternative constraint hypothesis was rejected as the optimal topology.

Molecular clock calibration

In order to examine and interpret patterns of speciation among the tragelaphid antelope, a timescale of lineage turnover is necessary. The molecular clock hypothesis, first suggested in the 1960s (Zuckerkandl and Pauling 1965) proposes that DNA sequences evolve at rates that are constant across lineages and time thus allowing divergences to be estimated in the absence of fossil evidence (Arbogast *et al.* 2002). Following the suggestion that members of the Bovidae accumulate mutations in a clock-like manner over time (Gatesy *et al.* 1997, Hassanin and Douzery 1999a, Matthee and Robinson 1999a, Matthee *et al.* 2001), molecular clocks for each DNA region may be justifiable. The ability of a molecular clock to accurately date lineage divergence is, however, limited by the assumption that the rate of evolution estimated for a lineage, and for a specific gene, is an accurate indication of rate of evolution in other lineages for the same gene. Lineage-specific variation has been well documented in a number of groups (DeBry 1992, Takezaki *et al.* 1995) and has been suggested to be particularly pronounced during times of rapid radiation, when normal evolutionary interactions are accelerated (Bromham *et al.* 2000, Adkins *et al.* 2001). Lineage rate variation between members of the tribe Tragelaphini within the nuclear and mitochondrial partitions was therefore tested in two ways. First, the substitution rates among lineages were compared relative to a closely related

outgroup using the relative rate test (RRTree 1.1.11, Robinson-Rechavi and Huchon 2000). Pairwise comparisons, initially including all ingroup taxa, were performed using *B. tragocamelus* as reference taxon. The choice of outgroup may effect the results of the relative rate test (Tajima 1993, Robinson *et al.* 1998) and for this reason additional analyses using *B. taurus* as reference taxon were also performed. Thereafter, taxa showing significant rate heterogeneity (0.01% confidence level) were re-examined following Rice's (1989) suggestion that inaccurate probability values could result from multiple testing.

As an alternative to the lineage specific relative rate test (and to explore the severity and influence of any rate heterogeneity detected by the relative rate test on the overall topology), a likelihood ratio based test was used to examine the rate of evolution among lineages. A maximum likelihood tree was constructed in PAUP* 4.0b10 using the model parameters suggested by MODELTEST 3.06. A second maximum likelihood tree was similarly constructed using the same model but with a molecular clock constraint. These two topologies were compared statistically using the S-H test with 100 replicates of RELL distribution. The lineage rate variation tests were performed on DNA regions characterized by complete taxonomic representation (i.e. cyt *b*, 16S rRNA, PRKCI, MGF, SPTBN, THY).

Even if a constant rate of change is assumed, divergence dates may vary as a result of taxon choice, DNA region, phylogenetic estimation method, correction factor used and calibration points (Benton 1999). Molecular clock estimates from single DNA regions may contain large statistical errors so corroboration from multiple markers may be a more reliable method of estimating divergence times (Kumar and Hedges 1998). In the absence of any significant rate differences among lineages, molecular clocks were calibrated independently for the four nuclear DNA regions (PRKCI, MGF, SPTBN and THY) and two mitochondrial DNA regions (cyt *b,* 16S rRNA) using the maximum likelihood corrected sequence divergence values plotted against dates derived from the fossil record. Maximum likelihood corrected sequence divergence values reduce the effect of among-site rate variation on molecular

clock estimates therefore facilitating comparisons among the different DNA regions investigated herein as well as published studies (Arbogast *et al.* 2002). The first calibration point was based on the divergence of the Bovini and Boselaphini (15 MYBP, Vrba 1979). The second calibration point marks the divergence between cattle (tribe Bovini) and *T. angasi,* reported to have occurred 11.5 MYBP (Hill *et al.* 1985). The split between the two lineages constituting the tribe Bovini, the cattle and buffalo (10 MYBP, Janecek *et al.* 1996) was used as the third calibration point. Following Kumar and Hedges (1998) the divergence time between any two groups was estimated using the average divergence time from all species pairs belonging to those groups. This procedure was performed for each DNA region separately and finally an average multiloci time estimate was calculated.

Chapter 3: Results and Discussion

Data description

The aligned supermatrix containing 10 DNA regions, 6 of nuclear origin and 4 of mitochondrial origin, contained 6000 characters, 2736 bp non-protein coding and 3264 bp protein coding sequences. The different DNA regions were characterized by variation in mode and rate of evolutionary change (Table III) but as was to be anticipated from the faster evolving markers, the combined mitochondrial partition contained a higher proportion of variable characters (28.9%). Interestingly, the low among-site variation (α = 1.023) present in the mitochondrial DNA data was caused primarily by cytochrome *b* data (α = 2.023), while the other mitochondrial markers contained more variation (α = 0.162-0.736). This surprisingly low among-site variation in the cytochrome *b* region has also been detected in the tribe Reduncini (Birungi and Arctander 2001). The nuclear partition had less variable characters (13.1%) with all nuclear DNA regions showing greater among-site rate variation (α = 0-0.505). Although Bull *et al.* (1993) have cautioned against combining data with different among-site variation, Sullivan (1996) suggested that the combined analysis of DNA regions with different evolutionary constraints under a homogeneous reconstruction model (i.e. parsimony with all characters weighted equally) may reveal signal present in some DNA regions which may have been obscured by noise when analyzed separately. In other words, a combined analysis may lead to the inclusion of the hidden support contained within some independent DNA regions which could significantly improve the robustness of clades in the supermatrix topology.

As typical of mammalian lineages (Irwin *et al.* 1991) the mitochondrial partition showed a bias against guanine (A = 32.5%, T = 25.5%, C = 26.8%, G = 15.2%). The nuclear partition (A = 28.9%, T = 32.2%, C = 19.6%, G = 19.3%) displayed a general bias against both guanine and cytosine.

The consistency (CI) and retention index (RI) values (Table IV) were generally lower in the mitochondrial partition (CI: 0.4861 to 0.7500, RI: 0.5889 to

Table III: Patterns of variability among the independent DNA regions, mitochondrial partition, nuclear partition and supermatrix.

DNA partition	Total bp	% Protein Coding	% Variable Characters	Gamma distribution (α value)	Nucleotide Frequencies			
					%A	%T	%C	%G
12S rRNA	591	0	15.6	0.736	36.1	22.7	24.7	16.5
16S rRNA	348	0	23.9	0.489	34.9	24.6	21.3	19.2
Cyt *b*	1140	100	35.8	2.023	31.9	25.9	28.9	13.4
Cox III	783	100	31.0	0.162	27.2	40.9	28.8	29.0
Cyp 19	194	0	6.70	<0.001	33.2	33.4	18.0	15.4
Lactoferrin	341	0	19.1	<0.001	19.2	24.1	29.2	27.5
MGF	674	0	14.1	0.499	32.3	34.6	18.1	15.0
PRKCI	667	11	9.60	<0.001	34.3	38.7	11.2	16.0
SPTBN	765	16	16.1	0.505	23.2	27.2	24.8	24.8
THY	667	34	9.90	0.139	30.1	32.6	19.3	18.1
Exon	2316	100	672	-	30.2	26.3	27.8	15.8
Intron	3684	0	562	-	30.8	30.7	20.3	18.3
Mitochondrial	2862	67.2	28.9	1.023	32.5	25.5	26.8	15.2
Nuclear	3138	12.8	13.1	0.318	28.9	32.2	19.6	19.3
Supermatrix	6000	38.6	20.6	0.606	30.7	28.9	23.1	17.3

Table IV: Summary of the topologies obtained from analysis of the independent DNA regions, mitochondrial partition, nuclear partition and supermatrix. The number of parsimony informative characters, number of trees, optimal tree length, consistency index (CI) and retention index (RI) values are given for each parsimony tree. Nodes A-G correspond to those in Figure IV, **X** indicates where a node was not retrieved, ? indicates where a nodes could not be retrieved due to missing taxa.

Data partition	Method	No. of pars. informative characters	No. of equally parsimonious trees	Tree length	CI	RI	Nodes A	B	C	D	E	F	G	No. nodes with >70% bootstrap or 95% posterior probability
12S rRNA	Pars.	51	94	142	0.5510	0.5882	51	X	X	X	X	X	X	0
	ML						51	X	X	X	X	X	X	0
	Bayesian						0.70	X	X	X	X	X	X	0
16S rRNA	Pars.	61	3	151	0.5547	0.7311	72	77	X	X	X	X	X	2
	ML						73	63	X	X	X	X	X	1
	Bayesian						0.99	1.00	X	X	X	X	X	2
Cyt *b*	Pars.	290	3	904	0.4861	0.6346	81	86	67	96	X	90	58	4
	ML						100	88	79	94	X	93	70	6
	Bayesian						1.00	0.99	1.00	1.00	X	1.00	0.79	5
CoxIII	Pars.	153	161	13	0.7500	0.8333	85	75	91	X	92	88	96	6
	ML						98	50	58	X	79	88	98	4
	Bayesian						1.00	0.84	0.94	X	1.00	0.97	1.00	4
Cyp19	Pars.	4	13	15	0.8000	0.8571	X	X	X	?	X	X	X	0
	ML						X	X	52	?	X	X	X	0
	Bayesian						X	X	X	?	X	X	X	0
Lactoferrin	Pars.	19	9	73	0.8889	0.9231	100	X	X	?	76	X	X	2
	ML						100	X	X	?	73	X	X	2
	Bayesian						1.00	X	1.00	?	0.99	X	X	3
MGF	Pars.	37	5	116	0.7547	0.8632	100	X	X	X	X	X	X	1
	ML						99	88	X	X	X	X	X	2
	Bayesian						1.00	X	X	X	X	X	X	1
PRKCI	Pars.	17	104	52	0.8261	0.9149	96	X	X	X	X	X	X	1
	ML						93	X	X	X	X	X	X	1
	Bayesian						1.00	X	X	X	X	X	0.64	1
SPTBN	Pars.	37	1	142	0.7736	0.8667	89	X	92	92	82	X	97	5
	ML						80	57	85	95	84	X	90	5
	Bayesian						1.00	X	0.97	1.00	1.00	X	1.00	5
THY	Pars.	31	24	75	0.7209	0.8400	100	X	X	X	X	X	X	1
	ML						100	X	X	X	X	X	X	1
	Bayesian						1.00	X	X	X	X	X	X	1
Mitochondrial	Pars.	550	1	1699	0.4753	0.5527	100	100	96	78	67	94	94	6
	ML						100	100	96	87	60	99	100	7
	Bayesian						1.00	1.00	1.00	1.00	0.54-0.57	1.00	1.00	6
Nuclear	Pars.	145	6	491	0.7162	0.8215	100	X	87	75	X	X	91	4
	ML						100	X	95	77	X	X	89	4
	Bayesian						1,00	X	1.00	1.00	X	X	1.00	4
Supermatrix	Pars.	700	2	2203	0.5043	0.5938	100	98	100	96	X	96	99	6
	ML						100	71	100	99	X	98	97	6
	Bayesian						1.00	0.96	1.00	1.00	0.75	1.00	1.00	6
% of times nodes were recovered							92.41	61.54	61.54	38.46	38.46	30.77	53.85	

0.8333) when compared to the nuclear partition (CI: 0.7209 to 0.8261, RI: 0.8400 to 0.9149). This may indicate that the characters contained in the mitochondrial partition are more prone to the effects of saturation. In particular, the cytochrome *b* DNA region with the lowest CI value (0.4861) may be especially vulnerable to saturation. Although saturated characters (e.g. third codon positions) are often removed from phylogenetic analysis, it has been suggested that not all homoplasic characters are real homoplasies and their retention may therefore increase phylogenetic structure (Källersjö *et al.* 1999). Consequently, all characters were included in analysis. The supermatrix's empirical TI/TV ratio increased from 5.43 to 9.01 when the mitochondrial partition was considered but decreased to 2.19 when the only the nuclear partition was examined. The TI/TV ratios of the supermatrix, mitochondrial partition and nuclear partition (excluding outgroups) were 6.75, 1.56 and 12.16 respectively.

Sequence divergence

Pairwise maximum likelihood sequence divergence values calculated for the supermatrix among tragelaphid lineages ranged from 2.45% (*T. euryceros* and *T. spekei*) to 7.88% (*T. imberbis* and *T. buxtoni*, Table V). Pairwise sequence divergence values among individuals of the same species ranged from 0.02% within the geographically isolated *T. buxtoni* lineage, to 0.69% within the geographically widespread *T. scriptus* lineage. Comparisons of outgroup and ingroup taxa revealed sequence divergence values of 8.55% (*B. tragocamelus* and *T. oryx*) to 17.87% (*A. melampus* and *T. scriptus*). The mitochondrial partition displayed higher sequence divergences of 4.84% (*T. euryceros* and *T. spekei*) to 14.61% (*T. imberbis* and *T. buxtoni*). The nuclear partition's values deviated slightly from those observed for the supermatrix and mitochondrial partitions with *T. derbianus* and *T. oryx* (0.60%) sharing the closest relationship, and *T.spekei* and *T. imberbis* the most distant relationship (2.07%).

Congruence of different data partitions

The partition homogeneity test based on incongruent length differences indicated that there was no significant conflict between the mitochondrial

Table V: Percentage sequence divergence values from the supermatrix among the tragelaphid species. Values below the diagonal represent the uncorrected P values, the values above the diagonal represent the GTR+G+I parameter corrected sequence divergence values.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
T. derbianus (1)	***	2.71	2.87	4.13	4.18	5.58	4.85	4.63	4.43	4.46	4.26	4.25	4.42	4.58	4.28	16.09	10.00	10.13	9.29
T. oryx (2)	2.40	***	0.29	3.33	3.40	6.31	5.66	4.63	4.91	4.28	4.47	3.42	4.66	4.59	3.75	16.20	11.65	11.04	8.55
T. oryx (3)	2.48	0.28	***	3.49	3.49	5.00	4.82	4.67	4.38	4.41	3.58	3.51	3.96	4.09	3.78	15.26	9.62	10.17	8.94
T. strepsiceros (4)	3.35	2.81	2.88	***	0.18	7.01	5.95	5.13	5.11	4.56	5.08	4.36	4.96	4.99	4.52	17.47	12.70	11.92	9.35
T. strepsiceros (5)	3.39	2.85	2.89	0.17	***	6.16	5.27	5.12	4.44	4.47	4.30	4.28	4.38	4.64	4.45	16.43	10.74	10.65	9.78
T. imberbis (6)	4.31	4.76	3.93	5.12	4.61	***	7.02	5.80	7.88	6.50	7.22	5.90	7.31	7.38	5.73	17.31	12.16	12.50	10.24
T. angasi (7)	3.81	4.36	3.79	4.49	4.07	5.12	***	0.29	7.20	6.03	6.29	5.12	6.56	6.03	4.81	17.52	12.42	12.25	9.80
T. angasi (8)	3.65	3.68	3.67	3.95	3.93	4.39	0.28	***	5.90	5.94	5.10	5.10	5.51	5.22	4.99	16.51	10.31	10.85	10.22
T. buxtoni (9)	3.56	3.92	3.53	4.02	3.57	5.59	5.26	4.45	***	0.02	4.52	3.53	4.84	5.00	4.02	17.63	13.19	12.78	10.60
T. buxtoni (10)	3.56	3.49	3.54	3.62	3.57	4.80	4.54	4.45	0.02	***	3.63	3.56	4.01	4.17	3.99	16.60	10.69	11.29	10.56
T. euryceros (11)	3.47	3.66	3.01	4.01	3.49	5.23	4.72	3.98	3.66	3.01	***	0.40	2.84	3.67	2.90	17.72	12.47	12.68	9.42
T. euryceros (12)	3.45	2.93	2.96	3.52	3.46	4.47	4.00	3.95	2.95	2.96	0.39	***	2.45	2.82	2.83	16.11	10.07	10.72	9.60
T. spekei (13)	3.56	3.77	3.25	3.92	3.50	5.26	4.85	4.19	3.85	3.25	2.45	2.15	***	3.89	3.22	17.86	13.20	12.84	9.53
T. scriptus (14)	3.67	3.72	3.37	3.94	3.70	5.34	4.56	4.04	3.98	3.39	3.06	2.46	3.21	***	0.69	17.87	12.61	12.61	9.57
T. scriptus (15)	3.45	3.14	3.13	3.60	3.54	4.36	3.78	3.87	3.27	3.24	2.49	2.44	2.71	0.65	***	15.33	9.89	10.52	9.39
A. melampus (16)	8.70	8.90	8.41	9.28	8.78	9.17	9.18	8.71	9.29	8.81	9.36	8.71	9.40	9.39	8.38	***	16.22	14.86	13.82
B. taurus (17)	6.48	7.21	6.28	7.62	6.80	7.37	7.46	6.60	7.84	6.77	7.51	6.48	7.80	7.55	6.40	9.04	***	9.80	8.76
S. caffer (18)	6.54	7.01	6.59	7.37	6.80	7.60	7.50	6.88	7.74	7.06	7.75	6.87	7.75	7.67	6.78	8.53	6.57	***	8.58
B. tragocamelus (19)	6.12	5.78	5.97	6.10	6.33	6.52	6.34	6.54	6.73	6.68	6.17	6.29	6.20	6.18	6.16	8.03	5.98	5.85	***

partition and nuclear partition ($P = 0.11$). The comparisons between the 0.95 Bayesian topologies with the maximum likelihood topologies for each of the 10 DNA regions analysed, revealed no conflict suggesting that each of the independent DNA regions had evolved under a similar, but not necessarily identical underlying topology (Buckley *et al.* 2002). Both partition homogeneity and the Bayesian approaches supported the combination of data into a supermatrix.

Phylogenetic estimation

The stability of nodes assessed by parsimony bootstrap and decay analysis of the mitochondrial partition, nuclear partition and supermatrix are presented in Figure II. Maximum likelihood bootstrap analysis using the LRT model parameters suggested by MODELTEST 3.06 (Table VI) resulted in topologies shown in Figure III. Differential weighting of the mitochondrial partition, nuclear partition and supermatrix did not result in improved resolution. Similarly, coding gaps as a 5^{th} character did not improve resolution of the mitochondrial partition, nuclear partition and nor the supermatrix. Consequently all characters were equally weighted and gaps coded as missing data. Only the topologies derived from these analyses are presented and discussed.

Bayesian analysis employing the different optimal models of evolution determined by MODELTEST 3.06 for each DNA region did not result in improved resolution for either of the three data partitions. The Bayesian inference was thus assumed to be independent of model choice. This is in contrast to the literature which often reflects increased support when different models are specified (Sullivan and Swofford 1997, Cunningham *et al.* 1998, Kelsey *et al.* 1999) but in our case the results are probably biased by the limited number of characters available for some independent DNA regions. The three independent Bayesian runs performed on the mitochondrial, nuclear and supermatrix partitions converged on identical topologies with similar log-likelihood scores. Graphic representation of the three runs indicate that convergence of the Markov chains was reached no later than 20 000 generations in all analysis. Consequently all topologies generated prior to this

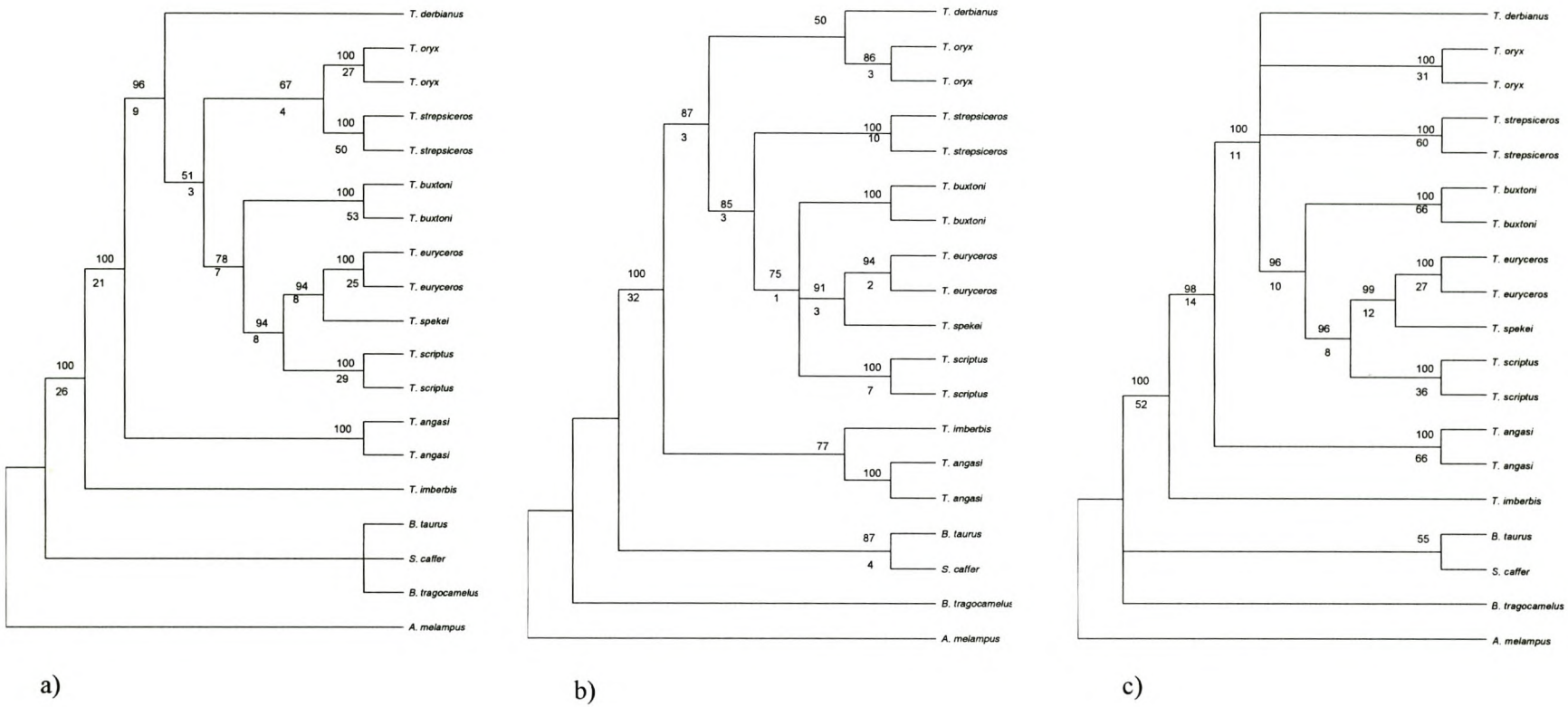


Figure II: Parsimony bootstrap topologies for a) mitochondrial partition, b) nuclear partition and c) supermatrix. The values above the branches indicate bootstrap support, the values below the branches indicate decay index values for the respective nodes. Tree lengths are indicated in Table IV.

Table VI: The model parameters selected for each DNA region as well as the mitochondrial partition, nuclear partition and supermatrix by MODELTEST 3.06.

		Rates of nucleotide change					
DNA partition	**LTR Model**	**A-C**	**A-G**	**A-T**	**C-G**	**C-T**	**G-T**
12S rRNA	TVM+I+G	0.221	41.29	6.054	0.823	41.29	1.000
16S rRNA	TrN+I+G	1.000	12.01	1.000	1.000	24.09	1.000
Cyt *b*	GTR+I+G	1.606	28.24	2.280	0.355	35.03	1.000
Cox III	TrN+G	1.000	29.87	1.000	1.000	24.80	1.000
Cyp 19	HKY	-	-	-	-	-	-
Lactoferrin	K80	-	-	-	-	-	-
MGF	K81uf+G	1.000	2.965	0.357	0.357	2.965	1.000
PRKCI	HKY	-	-	-	-	-	-
SPTBN	K80+G	-	-	-	-	-	-
THY	HKY+G	-	-	-	-	-	-
Mitochondrial	GTR+I+G	3.842	52.14	4.502	0.896	77.76	1.000
Nuclear	HKY+G	-	-	-	-	-	-
Supermatrix	GTR+I+G	1.498	9.933	1.064	0.719	15.73	1.000

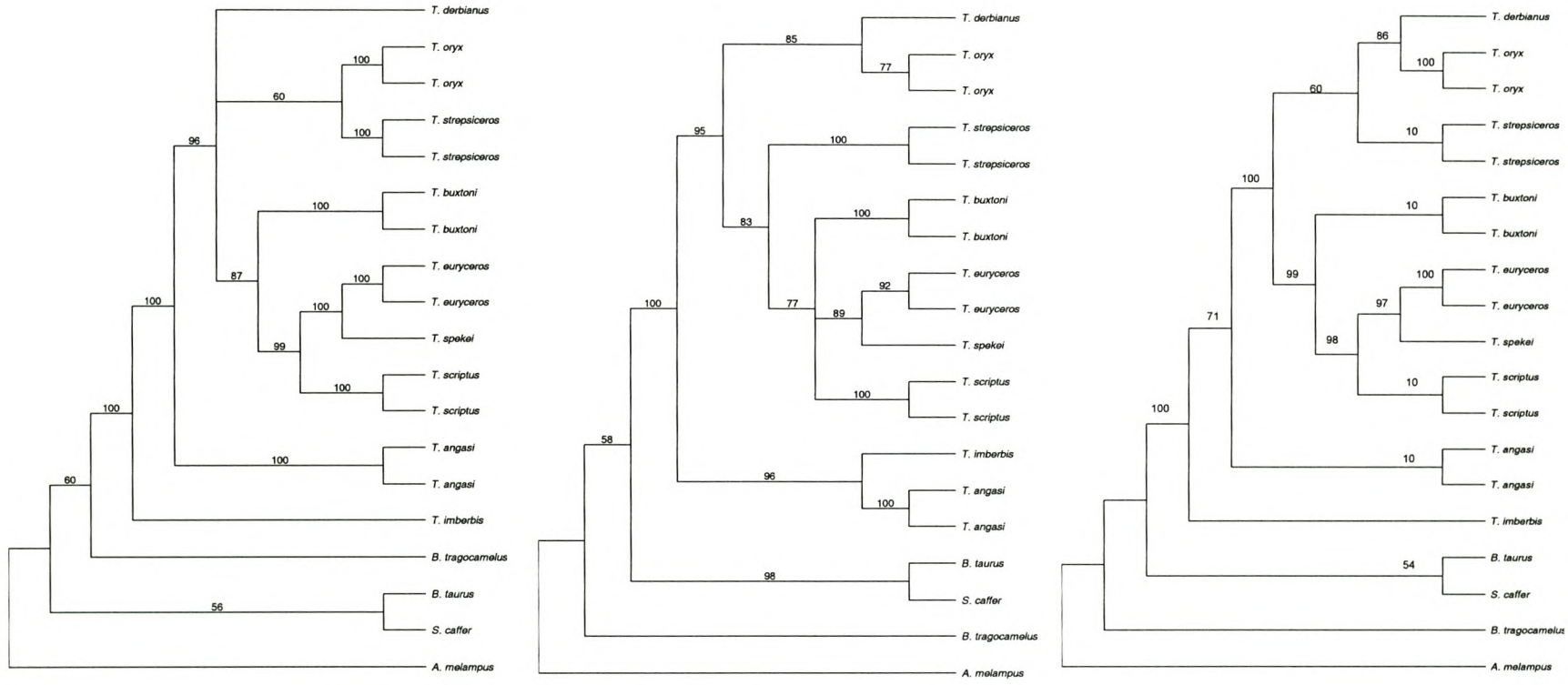


Figure III: Maximum likelihood bootstrap topologies for a) mitochondrial partition, b) nuclear partition and c) supermatrix. The values above the branches indicate bootstrap support for the respective nodes.

point were removed as "burnin". A 50% majority rule consensus tree based on single runs was constructed for each partition (Figure IV). The posterior probability values for each run were in general agreement. Slight differences in values were noted on the respective topologies (Figure IV). The 95% and 99% posterior intervals of mitochondrial and nuclear partitions and the supermatrix resulted in considerably fewer unique topologies relative to those occurring within the same posterior intervals of the individual DNA regions (Table VII). This decrease in topologies reflects the increase in resolution obtained from the combined analysis of DNA regions (Buckley *et al.* 2002).

In an attempt to compare how well each independent DNA region recovered the phylogeny, seven nodes of interest (A-G) were identified and these are marked on the Bayesian supermatrix topology (Figure IV). Nodes were selected for their ability to explain associations between the different tragelaphid species. The branch support values retrieved for selected nodes resulting from the analysis of the ten DNA regions as well as the combined analysis of the mitochondrial partition, nuclear partition and supermatrix under the various methods of phylogenetic estimation showed that several independent DNA regions did not perform well under most methods of analysis (Table IV). This includes 12S rRNA, cyp 19, cox III, lactoferrin, PRKCI, MGF and THY and is probably due to the limited number of nucleotide substitutions among taxa and/or the reduced taxonomic representation for some DNA regions (e.g. cox III, lactoferrin). The combined analyses of these DNA regions did, however, significantly improve resolution.

Parsimony bootstrap, maximum likelihood bootstrap and Bayesian posterior probability values resulting from the mitochondrial, nuclear and supermatrix topologies were compared graphically (Figure V). The points in the plots were not independent and cannot be assessed statistically but nonetheless provide an interesting comparison. Both the parsimony and maximum likelihood bootstrap values were highly correlated with the Bayesian inference posterior probability values. Although recent studies (Murphy *et al.* 2001, Suzuki 2002) have suggested that Bayesian posterior probability values may overestimate

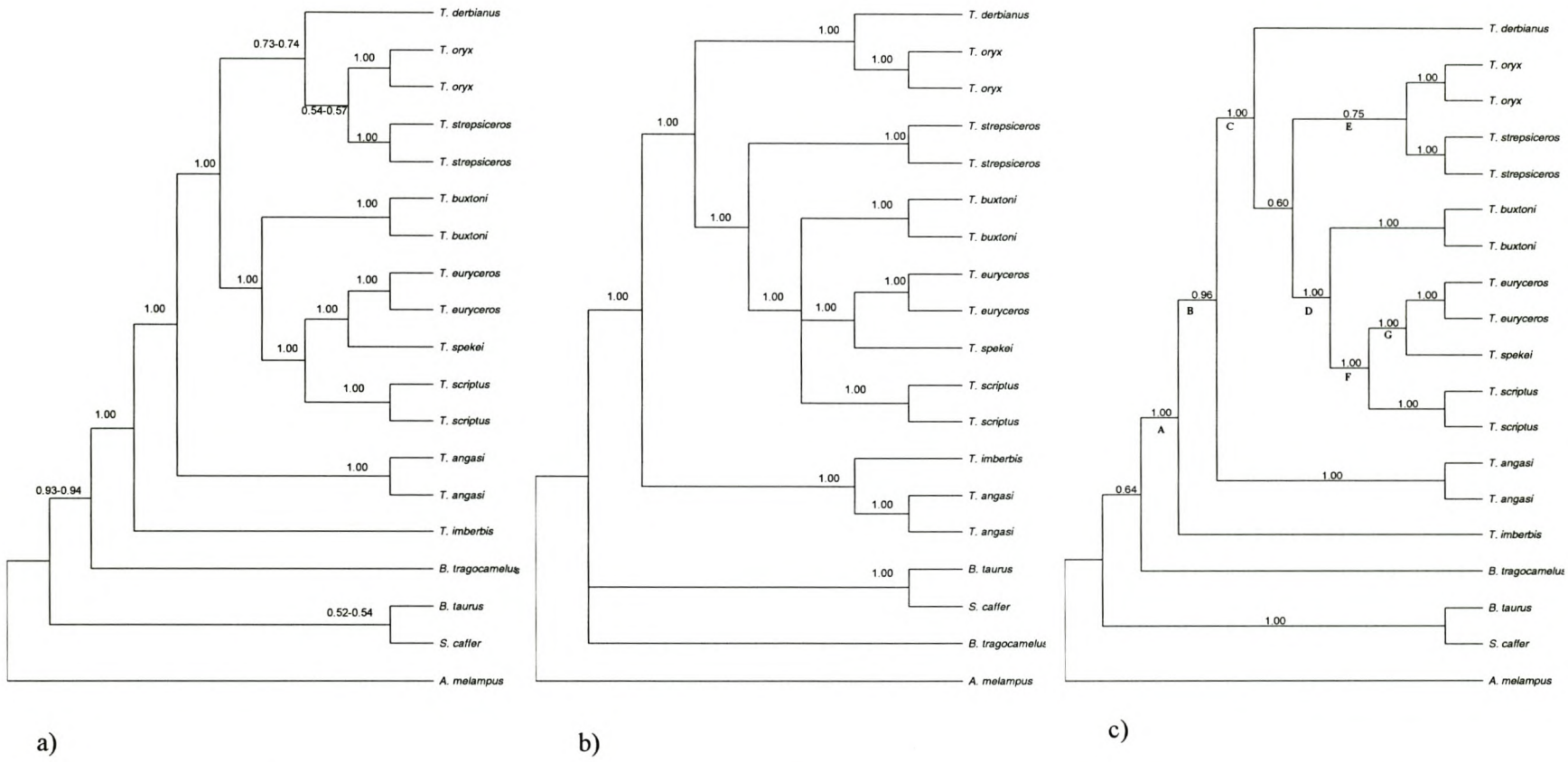


Figure IV: Bayesian topologies for a) mitochondrial partition, b) nuclear partition and c) supermatrix. The values above the branches indicate posterior probabilities for the respective nodes.

Table VII: The number of topologies found within the 50%, 90%, 95% and 99% posterior intervals of each DNA region, the mitochondrial partition, the nuclear partition and supermatrix.

	Posterior Intervals			
DNA Region	**50%**	**90%**	**95%**	**99%**
12S	7346	13346	14096	14696
16S	80	1265	2015	2615
Cox III	1	8	15	66
Cyp 19	7237	13237	13987	14587
Cyt *b*	19	131	212	439
Lactoferrin	78	238	278	362
PRKCI	6492	12492	13242	13842
MGF	1595	6549	7299	7899
SPTBN	22	304	517	994
THY	3328	9328	10078	10678
Mitochondrial	2	5	10	18
Nuclear	2	7	8	8
Supermatrix	1	8	11	16

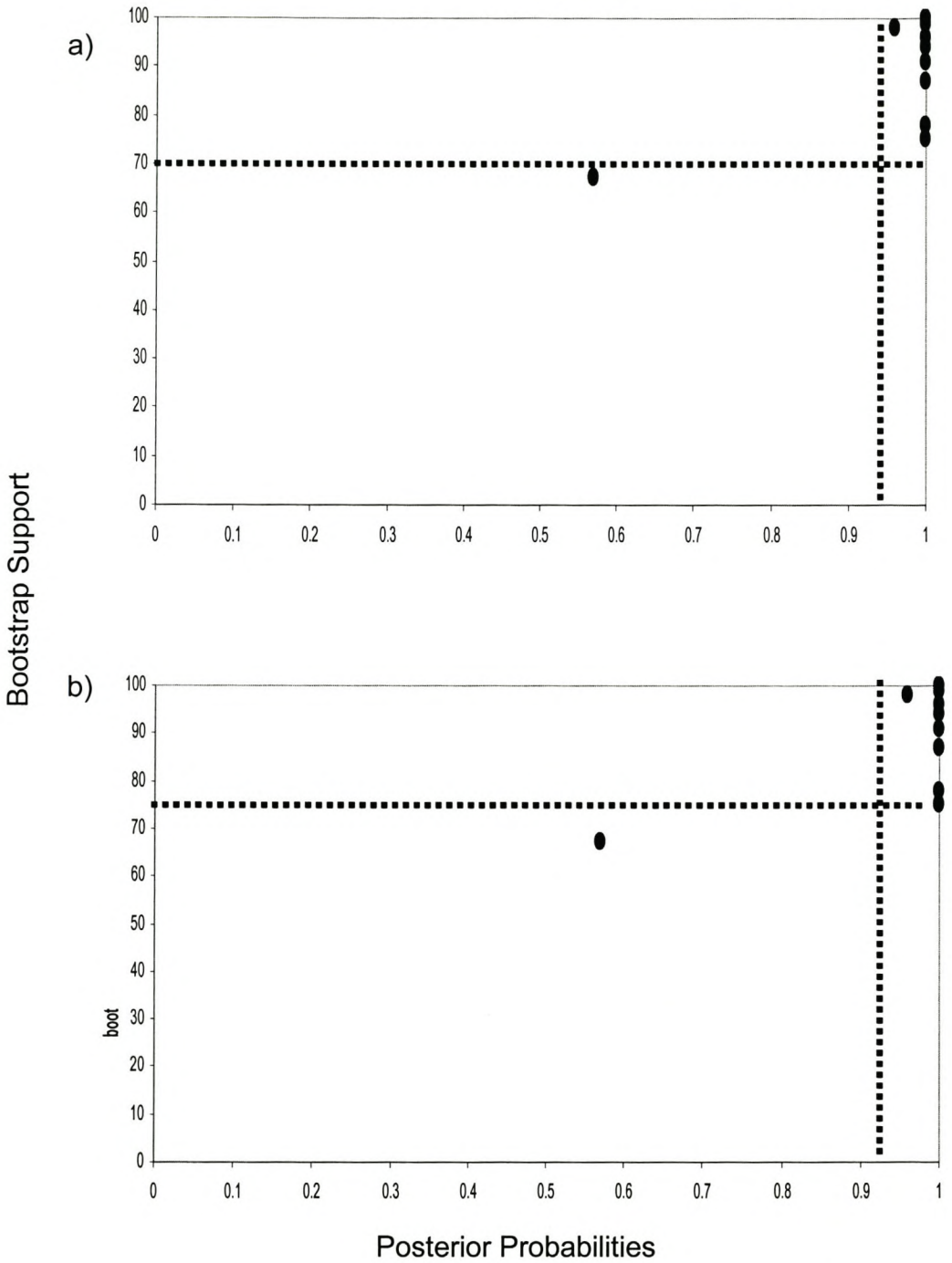


Figure V: Comparisons of Bayesian posterior probabilities with a) parsimony bootstrap support and b) maximum likelihood bootstrap support. Dotted lines represent where 70% bootstrap support corresponds to 0.95 posterior probability.

branch support, our findings support the assumption that bootstrap support of > 70% generally corresponds to 0.95 posterior probability (e.g. Leaché and Reeder 2002). For example, node E, although weakly supported by the parsimony and maximum likelihood bootstrap of the mitochondrial partition (67% parsimony, 60% maximum likelihood), was similarly weakly supported using Bayesian inference (0.57 posterior probability).

The tragelaphid phylogeny

The monophyletic status of the Tragelaphini tribe (node A) was confirmed by both the nuclear partition and mitochondrial partition under the various methods of analysis and it was therefore not surprising that this node was similarly supported by the supermatrix analysis (Table IV). The phylogenetic pattern retrieved by the analyses of the mitochondrial and nuclear and supermatrix data was found to be strongly influenced by the ecological separation of lineages and did not support several associations suggested by morphological studies.

Node C, confirming the basal position of *T. imberbis* and *T. angasi* was retrieved by all combined analyses irrespective of method used. A unique 31bp deletion (Appendix: 5459-5490) in the SPTBN DNA region characterized *T. euryceros*, *T. scriptus*, *T. buxtoni*, *T. spekei*, *T. oryx*, *T. derbianus* and *T. strepsiceros*. This indel was not present in *T. imberbis*, *T. angasi* nor the outgroup taxa, thus confirming the basal position of *T. imberbis* and *T. angasi* within the Tragelphini tribe. Although the nuclear topology suggested an association between *T. imberbis* and *T. angasi* (77% pars, 96% ML, 1.00 Bays), this was not confirmed by decay analysis.

Node D (which unites *T. buxtoni*, *T. euryceros*, *T. spekei* and *T. scriptus,* all species adapted to closed forest living), was consistently retrieved by both the mitochondrial and nuclear partitions under all methods of analysis. The monophyletic status of this group was confirmed by the analysis of the supermatrix showing high branch support (96% pars, 99% ML and 1.00 Bays) and decay values (optimal tree length of 2203 required 10 additional steps to collapse node D). The close evolutionary relationship between *T. spekei* and

T. euryceros suggested previously by mitochondrial data (Matthee and Robinson 1999a) was similarly retrieved by the extended mitochondrial DNA data set presented herein (94% pars, 100% ML and 1.00 Bays). Moreover, this association was confirmed by the nuclear partition as well as the analysis of the supermatrix under various methods of analysis. The faster evolving mitochondrial DNA yielded better resolution of the terminal nodes of this closed forest group, and placed *T. scriptus* as a sister lineage to the *T. spekei* and *T. euryceros* clade (94% pars, 99% ML, 1.00 Bays). This was similarly reflected in the supermatrix topology (node F, 96% pars, 98% ML, 1.00 Bays).

An association between *T. derbianus*, *T. oryx* and *T. strepsiceros,* which are united by shared physiological adaptations to an arid savannah environment was not consistently retrieved in all methods of analysis. The morphologically similar *T derbianus* and *T. oryx* were grouped as sister taxa by maximum likelihood analysis of the nuclear partition (85% bootstrap) and the supermatrix (86% bootstrap). Node E, which unites *T. strepsiceros* and *T. oryx* in the analysis of the mitochondrial partition and supermatrix was not supported by the nuclear topology which favored a close association between *T. strepsiceros* and the four species in the closed forest group (*T. buxtoni*, *T. euryceros*, *T. spekei* and *T. scriptus*). Branch support values of 85%, 83% and 1.00 were obtained respectively with parsimony, maximum likelihood and Bayesian inference.

Distribution of support among the different data partitions

Nodes A, B, C, F and G were constantly retrieved irrespective of the removal of any DNA region suggesting that their retrieval by supermatrix analysis was not influenced by any single DNA region. Only nodes A, D and G in the supermatrix were statistically supported by Templeton's test (Figure VI). This was not surprising given that these nodes were constantly retrieved (bootstrap values > 75% and posterior probability values > 0.95) in both the mitochondrial and nuclear topologies. The unambiguous synapomorphies derived from each DNA region and obtained for each selected node (Figure VI) showed nine of the ten independent regions contained unambiguous

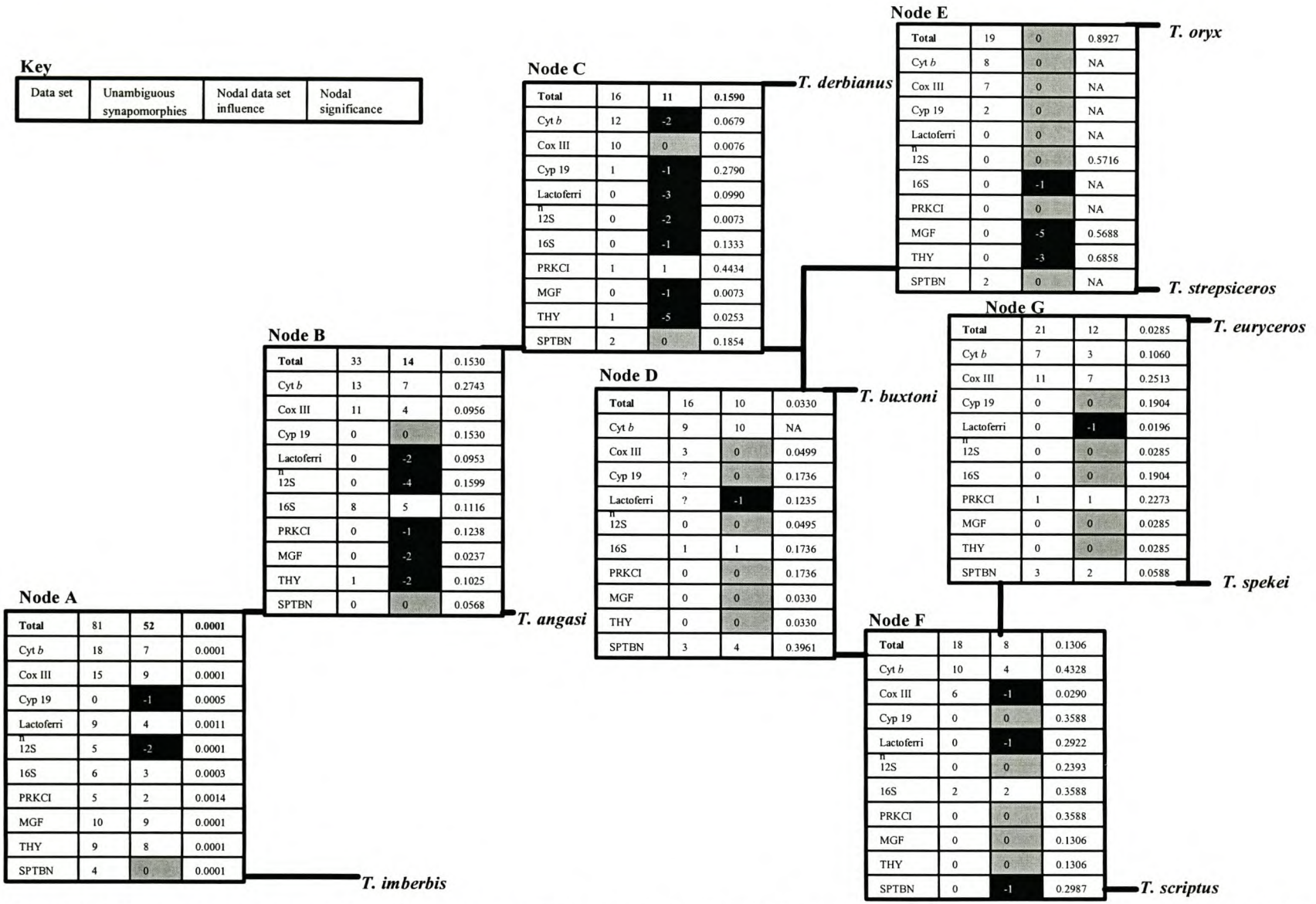


Figure VI: The number of unambiguous synapmorphies, nodal data set influence and nodal significance values for each DNA region at nodes A-G.

synapomorphies in support of the monophyletic status of the Tragelaphini tribe (node A).

The analysis of each DNA region independently resulted largely in unresolved topologies (Table IV). Surprisingly, however, many of these DNA regions provided positive character support for nodes when analyzed as part of the supermatrix (e.g. PRKCI for node C and G, as well as 16S rRNA for node D and F). These positive NDI values were supported by Templeton's test suggesting that their removal resulted in non-significant values in the truncated supermatrix analysis.

The monophyletic status of the tribe Tragelaphini (node A) was consistently well supported in all combined topologies (100% bootstrap support and 1.00 posterior probabilities). Moreover, it was significantly supported by Templeton's test although the NDI values suggest that the inclusion of aromatase cytochrome p-450 (cyp 19) and 12S rRNA within the supermatrix negatively effected this node. These and other marginally negative NDI values obtained for several of the nodes that were well supported by the combined analysis and the Templeton's tests (i.e. node D, F and G) may, however, result from the reduced numbers of variable characters (or the missing data) characterizing a number of DNA regions.

The distribution of probability values retrieved by Templeton's tests illustrate that different DNA regions are better able to resolve different nodes. For example, the removal of only five DNA regions resulted in non-significant *P*-values at node G (i.e. node G was unaffected by the removal of lactoferrin, 12S rRNA, MGF, THY and SPTBN) suggesting that the phylogentic signal in support of this node was contained within the cytochrome b, cox III, cyp 19, 16s rRNA and PRKCI data sets.

Node D which was well supported by cytochrome *b* sequences (NDI = 10), was not resistant to the removal of the cytochome *b* data set and thus the significance of this data set at that particular node could not be calculated. Lee and Hugall (2003) suggested that larger data sets, such as our

cytochrome *b* data set (which contributed 41% of informative characters to the supermatrix), may generate inflated NDI scores and non-significant values when assessed using Templeton's test. The possibility exists that the *T. buxtoni* lineage diverged from the other closed forest species nearly contemporaneously, resulting in this relationship being supported only by the most rapidly evolving DNA markers. This would explain why both 16S rRNA and SPTBN, which contain relatively more variable characters that the other markers, retrieved positive NDI values at this node. Interestingly, the removal of these DNA regions from the supermatrix resulted in not significantly support for this node.

Tests of alternative molecular phylogenies

Our constraint analyses, which tested the robustness of some phylogenetic associations retrieved in the mitochondrial and supermatrix topologies (Table VIII), resulted in the rejection of a single origin for the species adapted to the drier savannah environment (*T. derbianus*, *T. oryx* and *T. strepsiceros, P* = 0.001). In contrast, this association could not be rejected by the nuclear partition (*P* = 0.193). The association between *T. derbianus* and *T. oryx* as suggested by morphological studies and supported in the nuclear topology was rejected by the analyses of the mitochondrial and supermatrix data sets (*P* = 0.001). By enforcing an association between *T. oryx* and *T. strepsiceros* in the nuclear tree a better topology was obtained than that suggested in the unconstraint tree and this implies that the association between these two species could not be rejected. An association between the two basal species, *T. angasi* and *T. imberbis*, suggested in the nuclear tree was rejected by both the mitochondrial and supermatrix topologies (*P* = 0.001).

Mitochondrial versus nuclear DNA markers

The interpretations of tragelaphid evolution has been characterized by conflicting morphological and mitochondrial DNA evidence (Gatesy *et al.* 1997, Hassanin and Douzery 1999a, Hassanin and Douzery 1999b, Matthee and Robinson 1999a). By including nuclear DNA sequence data we attempted to confirm or invalidate the relationships suggested by the morphological and mitochondrial DNA studies. The nuclear DNA regions

Table VIII: The statistical support for the alternative hypotheses on the tragelaphid phylogenetic relationships based on the Shimodaira-Hasegawa (1999) test.

Tree	**-inL**	**Δ-inL**	***P* value**
Mitochondrail and supermatrix unconstraint tree	19171.25		
Monophyletic savannah group	19231.15	59.89	0.001
T. derbianus sister to *T. oryx*	19247.82	76.57	0.000
T. derbianus sister to *T. strepsiceros*	19369.76	198.51	0.000
T. angasi sister to *T. imberbis*	19333.66	162.41	0.000
Nuclear unconstraint tree	19175.92		
Monophyletic savannah group	19185.34	9.42	0.193
T. oryx sister to *T. strepsiceros*	19168.19	7.73	0.165

were, as expected, considerably less variable than those of the mitochondrial partition. This lack of variability did not, however, adversely effect the phylogenetic analyses that utilized both mitochondrial as well as nuclear markers and the combined analysis yielded a robust and well resolved phylogeny for the tribe. Our study clearly demonstrates the ability of the nuclear intron sequences to resolve associations at taxonomic levels previously considered intractable using slower evolving nuclear markers (Macnish *et al.* 2002, Cortés-Ortiz *et al.* 2003). The higher CI and RI values generated by the nuclear DNA suggests that these markers were less prone to homoplasy and saturation in our study, which is often characteristic of mitochondrial DNA markers (e.g. Allard *et al.* 1992, Gatesy *et al.* 1997). On the other hand it is important to realize that at this taxonomic level, the nuclear DNA sequence data tend to be more useful only when a large number of characters are employed (i.e. the combined nuclear DNA data set presented herein). From these findings it is thus difficult to make recommendations for future research as to which markers will be most useful (mitochondrial or nuclear) at this taxonomic level. Clearly the different markers employed in the present study contribute phylogenetic resolution at different nodes and it is thus not unexpected that the combined approach generally outperformed any single gene analysis.

Tragelaphid systematics: molecules versus morphology

The delimitation of species on ecological grounds in the present study through the analysis of the mitochondrial and nuclear partitions both singularly as well as in combination supports the phylogeny presented herein. Based on DNA sequence data our findings suggest that the morphological characters previously used to describe the relationships within the Tragelaphini tribe are convergent and therefore not good indicators of phylogenetic relatedness. Such morphological convergence is not a new phenomenon and is prevalent in numerous mammalian orders (Chevret *et al.* 1993), specifically it has been observed within a number of bovid tribes (Estes 1991, Gatesy *et al.* 1997, Matthee and Robinson 1999a).

Given that our study is presently the most taxonomically comprehensive, for the Tragelaphini we consider our supermatrix phylogeny stable. In particular the monophyletic status of the tribe and the group containing the four closed forest species (*T. buxtoni*, *T. euryceros*, *T. spekei* and *T. scriptus*) as well as the basal position of *T. imberbis* and *T. angasi* which were resistant to bootstrap and DNA region removal.

Although our nuclear data are unable to confirm the association between *T. strepsiceros* and *T. oryx* suggested by the mitochondrial DNA, the constraint analysis which enforced an association between these two species could not be rejected as an alternative topology. A likely explanation may be that the nuclear DNA markers have mutational rates that are too slow to adequately resolve this relationship (i.e. not enough nucleotide changes have accumulated since common ancestry). Similarly, an association between the two most basal species *T. imberbis* and *T. angasi,* which was resolved by the more conservative nuclear markers, was not detected by the faster evolving mitochondrial DNA markers. The nuclear sequence data included in our study suggest the inclusion of all extant species into a single genus, *Tragelaphus*, confirming the conclusions of previous studies based on fossil evidence (Vrba 1987), mitochondrial DNA sequences (Essop *et al.* 1997, Gatesy *et al.* 1997, Hassanin and Douzery 1999a, Matthee and Robinson 1999a) and allozymes (Geogiadis *et al.* 1990).

The supermatrix approach

Phylogenetic analyses of the independent DNA regions produced largely unresolved topologies (Table IV), with only the highly variable cytochrome *b* data set resulting in a topology comparable to that obtained by analysis of the combined data (retrieving all nodes except node E). The simultaneous analysis of all mitochondrial markers, the nuclear markers and the subsequent merging of all sequence data into a single data matrix or "supermatrix" resulted in increased resolution and nodal support for both the deepest and more recent nodes of the tragelaphid evolutionary tree. This study therefore confirms Sullivan's (1996) suggestion that combining DNA regions with different evolutionary constraints leads to the incorporation of phylogenetic

signal present in some DNA regions which is often obscured in independent analysis by noise or weakened by the lack of sufficient phylogenetically informative substitutions. This was clearly evidenced by the NDI and Templeton's tests which show that unresolved independent DNA regions may often provide positive character support in supermatrix analysis even when the addition of some DNA regions (i.e. cox III, cyp 19 and lactoferrin) resulted in increased missing data. Additionally our supermatrix summarizes all currently available sequence data and, identifies areas of missing data for future research (for example, the tentative relationships suggested between *T. strepsiceros* and *T. oryx*, and *T. angasi* and *T. imberbis* by our sequence data will need further investigation).

Molecular clock

i) Among-lineage rate variation

The relative rate test (Robinson-Rechavi and Huchon 2000) performed using *B. tragocamelus* as reference taxon indicated no significant rate heterogeneity among lineages in the nuclear partition. Examination of the mitochondrial partition identified a single significant difference in substitution rates between *T. oryx* and *T. angasi* (P = 0.008). The relative rate test using *B. taurus* as outgroup species resulted in no significant difference in rate heterogeneity among lineages in the mitochondrial partition. Surprisingly, examination of the nuclear partition with *B. taurus* as reference taxon indicated a significant difference in evolution rate between the *T. buxtoni* and *T. scriptus* lineages (P = 0.008). A molecular clock could not be rejected by the S-H test for neither the nuclear (P = 0.202) nor the mitochondrial (P = 1.00) partitions. Given the fairly uniform rate of evolution among tragelaphid lineages suggested by both the relative rate test and the likelihood ratio based test, the application of a molecular clock for the nuclear partition and mitochondrial partition was deemed appropriate.

ii) Molecular dating

The potential pitfalls associated with using a single DNA region for molecular clock calibrations were addressed by using independent calibrations for the

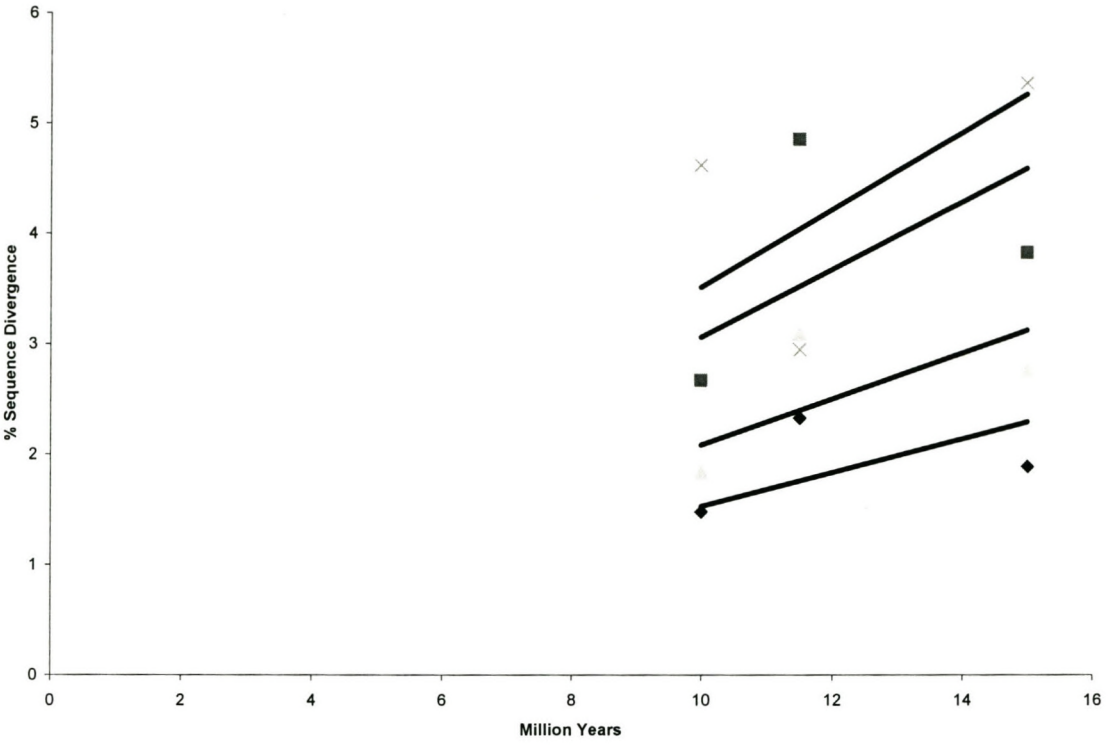


Figure VII a): Molecular clock regression analysis on SPTBN (×), MGF (■), THY (▲) and PRKCI (♦) based on corrected sequence divergence and fossil dating.

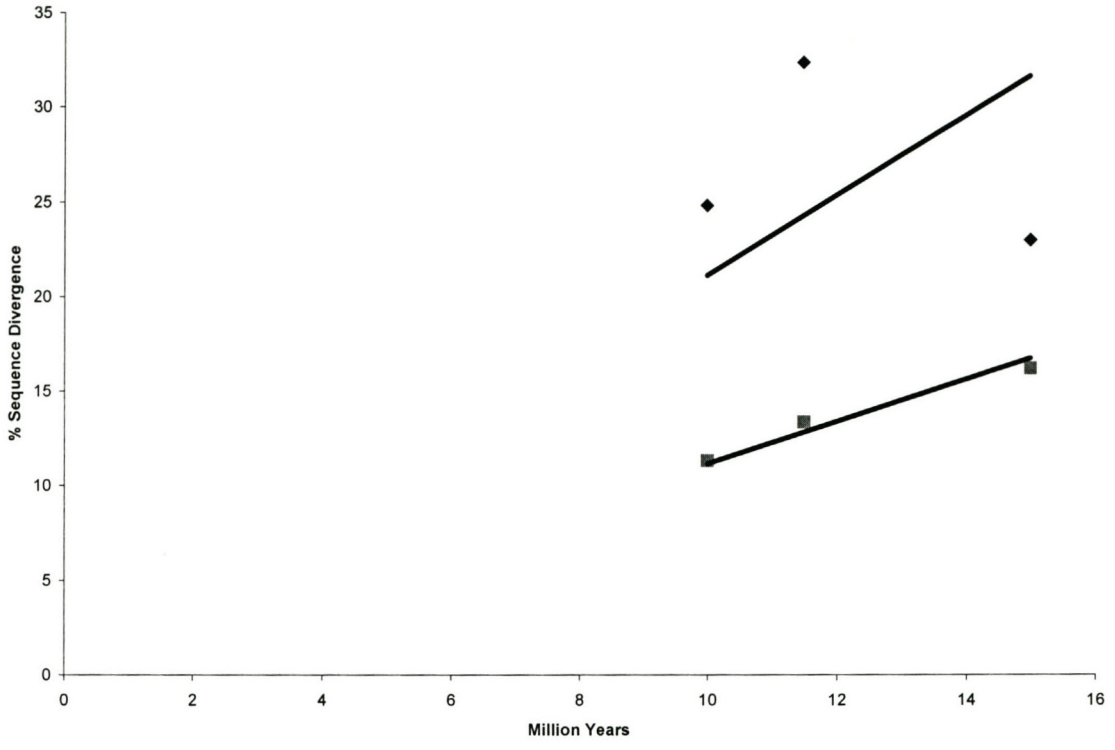


Figure VII b): Molecular clock regression analysis on cytochrome *b* (♦) and 16S rRNA (■) based on corrected sequence divergence and fossil dating.

mitochondrial and nuclear partitions. We followed Allard *et al.* 1992, Hillis *et al.* 1996, Matthee and Robinson 1999a and Flagstad *et al.* 2000 by applying regression analysis to the fossil dates (in millions of years before present) and sequence divergence values (Figure VII). To facilitate comparisons among the different DNA regions, maximum likelihood corrected sequence divergence values were calculated (by applying models selected in MODELTEST) and used in the regression analysis.

As anticipated, the nuclear markers were found to evolve more slowly than the mitochondrial DNA loci. Of the markers, the PRKCI partition was the most conserved DNA region (mutational rate of 0.153% per million years) and cytochrome *b* the most rapidly evolving. In the case of the latter the mutational rate was 8.29 times greater than the average mutational rate of the slower nuclear partition (0.2548% per million years, Table IX). Following Matthee and Robinson's (1999a) suggestion that transversional changes may increase linearly over time we decided on an additional molecular clock calibration using only cytochrome *b* transversion based sequence divergence values. The mutational rate was found to be 0.1995% transversions per million years. Although higher than the rate suggested for cytochrome *b* by Hassanin and Douzery (1999a, 0.13% transversions per million years) for the family Bovidae, our rate estimation within the tribe corresponds closely with that of Matthee and Robinson (1999a, 0.22% transversions per million years).

We place the divergence of the Tragelaphini at 13-19.4 MYBP (range of values refer to the extreme estimates obtained from the different DNA regions, Table X) based on clock calibrations from the four nuclear DNA regions. The split between the basal species *T. angasi* and *T. imberbis* (both species occurring in bushland habitats) and the other tragelaphid antelope was estimated to have occurred 5.3-6.7 MYBP. This was followed by a period of rapid ecological specialization resulting in those species that are confined to moist forest environments (*T. buxtoni*, *T. euryceros*, *T. spekei* and *T. scriptus*) and those adapted to more arid savannah environment (*T. derbianus*, *T. oryx* and *T. strepsiceros*); our estimates place this at 4.1-5.7 MYBP. Our second approach to date the cladogenesis within the Tragelaphini using the

Table IX: Details of the regression analyses and the evolutionary rates of the each DNA region used in the molecular clock calibrations.

DNA Region	R^2	Evolutionary rate per million years
Nuclear Partition		
PRKCI	0.861	0.15%
MGF	-0.053	0.31%
SPTBN	0.210	0.35%
THY	0.197	0.21%
Mitochondrial Partition		
cyt *b*	-2.127	2.11%
16S rRNA	0.948	1.12%

Table X: Summary of the divergence times (in millions of years before present) estimates using the molecular clocks.

Divergence event	Divergence times (MYBP)								
	PRKCI	MGF	THY	SPTBN	16S rRNA	Cyt *b*	Average nuclear	Average mitochondrial	Average multigene
Tragelaphini tribe diverge from other bovid tribes	16.94	15.69	19.42	13.09	13.79	15.07	16.28 (± 3.17)	14.43 (± 0.64)	15.67 (± 3.17)
T. imberbis and *T. angasi* diverge from the other Tragelaphids	6.65	5.74	5.26	6.72	9.72	8.69	6.09 (± 0.73)	9.21 (± 0.52)	7.13 (± 2.23)
Closed forest group diverge from the *T. derbianus*, *T. oryx* and *T. strepsiceros*.	5.74	5.64	4.14	5.35	5.93	5.79	5.22 (± 0.80)	5.86 (± 0.07)	5.43 (± 0.89)

mitochondrial DNA based calibration (Table X) corresponds well with those obtained from the nuclear DNA regions. Analysis of this marker places the initial tribal diversification within the family Bovidae at 13.8-15 MYBP (*cf* 13-19.4 MYBP based on nuclear sequences, see above). The split between the basal bushland species (*T. angasi* and *T. imberbis*) and the other tragelaphids was estimated to have occurred 8.7-9.7 MYBP (*cf* 5.3-6.7 MYBP based on nuclear sequences), with the split between the moist forest adapted species (*T. buxtoni*, *T. euryceros*, *T. spekei* and *T. scriptus*) and savannah adapted species (*T. derbianus*, *T. oryx* and *T. strepsiceros*) at approximately 5.8-5.9 MYBP (*cf* 4.1-5.7 MYBP based on nuclear sequences). Our divergence dates are broadly similar to estimates from previous molecular clock calibrations that suggest that the tribe emerged 14-16 MYBP, with later intertribal lineage turnover events occurring in the late Miocene or at the Miocene-Pliocene boundary (Vrba 1985a, Geogiadis *et al.* 1990, Hassanin and Douzery 1999a). The interspecific and intraspecific (Haltenorth 1963) variability in many of the morphological features previously used to delimit the group and the effects of the morphological convergence makes the tracking of species turn-over from the fossil record difficult. For example Gentry (1978) suggested that *T. strepsiceros* and *T. imberbis* diverged approximately 3 MYBP, which is clearly in conflict with the molecular phylogeny presented herein.

iii) Divergence estimates using multiple loci

Estimates from single genes are often plagued by large statistical errors (Kumar and Hedges 1998). The use of multilocus molecular clocks are thought to provide a more reliable estimate of lineage divergence, the underlying rational being that different DNA regions provide phylogenetic signal at different hierarchical levels. In combination therefore, they may provide the best overall resolution of a phylogeny (Hillis 1987). The six selected DNA regions were compared graphically (Figure VIII) in order to explore the accuracy of the dating across several DNA regions. Divergence estimates (in millions of years) are listed (Table IX) from the most conservative DNA region on the left (PRKCI- average evolutionary rate of 0.153 mutations per million years) to the most rapidly evolving on the right (cytochrome *b*- average evolutionary rate of 2.112 mutations per million years,

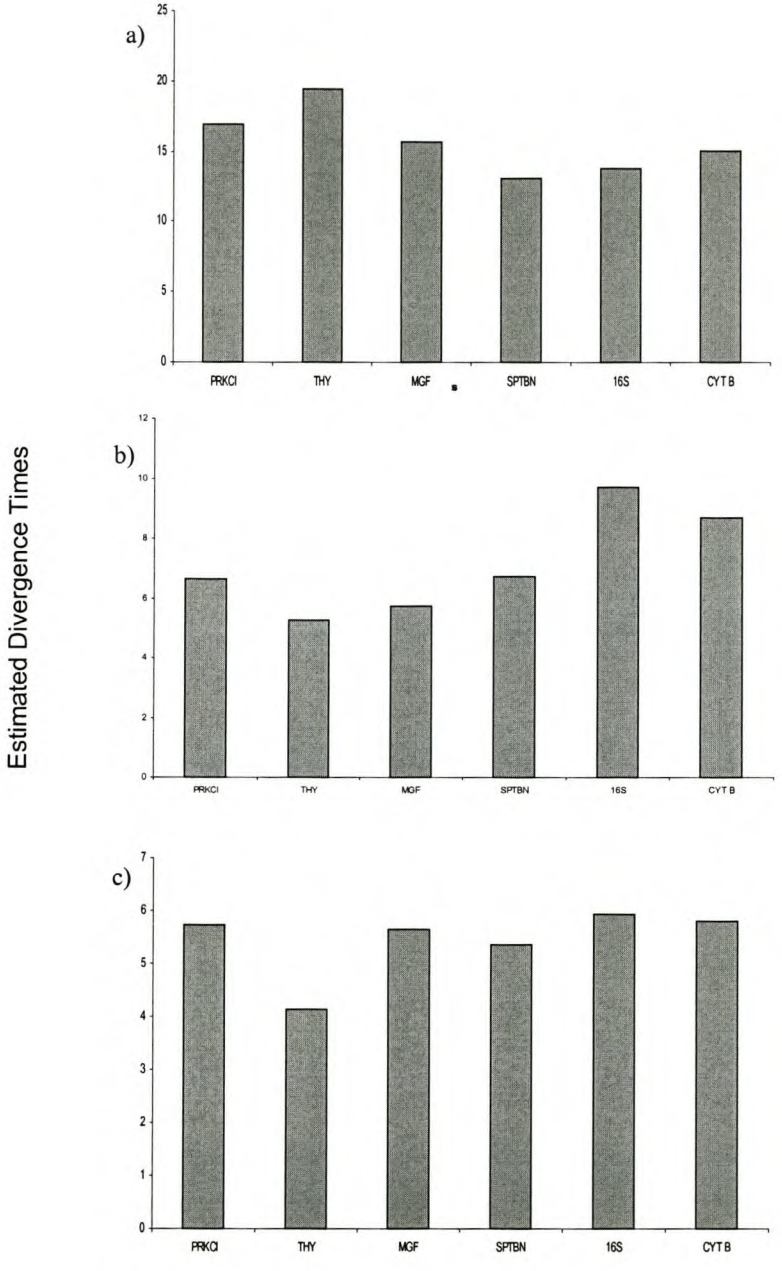


Figure VIII: The distributions of estimated divergence times for a) divergence of members of the Tragelphini from the outgroup taxa, b) the divergence of the two basal species (*T. imberbis* and *T. angasi*) from the remaining tragelaphid species and c) the split between the closed forest group (*T. buxtoni, T. euryceros, T. spekei* and *T. scriptus*) and the savannah adapted species (*T. derbianus, T. oryx* and *T. strepsiceros*).

Table IX). The distribution is skewed to the left when the oldest divergence event is examined (Figure VIII a) shifting to the right when estimates for the more recent divergence events are plotted (Figure VIII b and c). Our data suggest that the deeper events are generally more accurately estimated when using the slower evolving nuclear DNA regions (i.e. PRKCI and THY), while the faster rate of the mitochondrial DNA regions (i.e. 16S rRNA and cytochrome *b*) are more suitable for estimating more recent divergence events. The lack of a well-defined peak of distribution in Figure VIII c, suggests, however, that the most recent divergence event (the split between the closed forest and savannah species), is predicted equally well by both the slower and faster evolving DNA regions.

Correlation of phylogeny with Paleoclimatic change

Although the tempo governing tragelaphid evolution can be estimated using a molecular clock, the mode of speciation can only be examined by comparing the tragelaphid phylogeny with paleoclimatic changes in Africa. The present study is the first to include representatives of all nine recognized species and meets the criteria proposed by Barraclough and Nee (2001), who hold that patterns of speciation can only be accurately illuminated if all extant species from a selected group are sampled.

A number of speciation hypothesis have been proposed that acknowledge the role of climate change on lineage turnover. These include gradualist environmentalism, punctualist enviromentalism and the turnover-pulse hypothesis (Vrba 1985b). Gradualism, first proposed by Darwin, has natural selection as its foundation and emphasizes the interaction between organisms and the environment. This theory acknowledges the role played by climate change on lineage turnover but suggests that this may not be the only cause of speciation and extinction (Vrba 1980, Vrba 1984). Over time a constant "background noise" of turnover events are to be expected with an acceleration, or peak, in response to extremes in environmental change (Vrba 1980, Vrba 1984). The punctuated equilibrium hypothesis proposed by Eldredge and Gould (1972) suggests rapid genetic changes within geographically isolated populations as the main force driving lineage turnover.

Populations faced with new environmental pressures would be subject to directional selection (Vrba 1984). In small populations the limited number of individuals would allow for more rapid fixation of neutral alleles by random drift (Vrba 1984). The turnover pulse hypothesis (Vrba 1985a) suggests that speciation would not occur unless forced, or initiated, by changes in the physical environment. Speciation would occur in pulses over time, as predicted by the punctuated equilibrium hypothesis, and that changes in the physical environment would be captured in the fossil record and reflected across diverse phylogenies.

During the Miocene-Pliocene period Africa was subject to periods of climatic change (Cerling *et al.* 1997, Vrba 1995, Zachos *et al.* 2001). There was a general global cooling (associated with the development of the polar caps) with global temperatures declining sharply during both the mid-Miocene (approximately 14MYBP) and Pliocene (approximately 5MYBP, Cerling *et al.* 1997, Zachos *et al.* 2001). The decline in temperature during the mid-Miocene saw an increase in seasonality leading to the expansion of savannah grasslands and the contraction of humid forest zones into geographically isolated forest "islands" surrounded by drier bushland and savannah (Lindsay 1998). Such climatic change would have created conditions favorable for the evolution of drought-resistant species able to exploit the expanding grassland vegetation (Georgiadis *et al.* 1990). Cerling *et al.* (1997) suggested that during this period, grazing antelopes and hippos replaced chevrotains and anthracothers as the dominant artiodactyls. This evolutionary pattern is indeed reflected in the phylogeny of the Tragelaphini tribe presented herein. The times of sudden declines in temperature correspond well to times of rapid lineage turnover as estimated by both molecular clocks (Table X). The tribe Tragelaphini diverged from the other bovids during the mid-Miocene at approximately 15.7 MYBP, probably prompted by the cooling trend which reached a peak at approximately 14 MYBP. The species *T. angasi* and *T. imberbis* (representing the most basal species in the molecular phylogeny) separated from the other tragelaphids at approximately 7.1 MYBP. This initial divergence was followed by the ecological specialization of species with a split occurring between species adapted to a more tropical environment including

T. buxtoni, *T. euryceros*, *T. spekei* and *T. scriptus*, and species adapted to a more arid savannah environment including *T. derbianus*, *T. oryx* and *T. strepsiceros*. This final cladogenesis probably coincided with the Miocene-Pliocene boundary approximately 5.4 MYBP and the decline in global temperature approximately 5 MYBP (Cerling *et al.* 1997, Zachos *et al.* 2001). This correlation suggests that environmental change, and the associated increase in habitat diversity probably, was the major driver of cladogenesis within the Tragalephini. As predicted by Vrba's (1985a) turnover pulse speciation hypothesis, this same pattern of lineage turnover suggested for the tribe Tragelaphini is also reflected in other non-related taxa (Lacertidae, Fu 2000), which similarly underwent a rapid radiation into grassland habitats during the Miocene-Pliocene period.

In conclusion, the key to predicting what future influence climatic change will have on contemporary species lies in understanding the evolutionary mechanisms of the past. The fossil record suggests that long-term global climatic change leads to an alteration in the availability of appropriate habitats thereby accelerating rates of speciation and extinction in many groups (Vrba 1995). By matching the evolutionary events predicted by our phylogeny with periods of significant climatic change, we suggest that climatic change acted as a major evolutionary determinant within the tribe Tragelaphini. One cannot help but wonder what effect future climate change will have on the tragelaphids and, in a broader sense, biodiversity in Africa, given the impacts of human development, the fragmentation of habitats and the disruption of historic dispersal routes.

References

Adkins, R. M., Gelke, E. L., Rowe, D. & Honeycutt, R. L. (2001). Molecular phylogeny and divergence time estimates for major rodent groups: Evidence from multiple genes. *Mol. Biol. Evol.* **18**: 777-791.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Trans. Autom. Contr.* **19**: 716-723.

Alden, P. C., Estes, R. D., Schlitter, D. & MacBride, B. (1995). National Audubon Society Field Guide to African wildlife. Knopf, New York.

Allard, M., Miyamoto, M., Jarecki, L., Kraus, F. & Tennant, M. (1992). DNA systematics and evolution of the artiodactyls family Bovidae. *Proc. Natl. Acad. Sci. USA.* **89**: 3972-3976.

Ansell, W. F. H. (1971). Order Artiodactyla. In: *The Mammals of Africa: an identification manual.* (eds) J. Meester & H. W. Setzer. Smithsonian Institute Press, Washington.

Arbogast, B. S., Edwards, S. V., Wakeley, J., Beerli, P. & Slowinski, J. B. (2002). Estimating divergence times from molecular data on phylogenetic and population genetic timescales. *Annu. Rev. Ecol. Syst.* **33**: 707-740.

Atchley, W. R. & Fitch, W. M. (1991). Gene trees and the origins of inbred strains of mice. *Science* **254**: 554-558.

Barraclough, T. G. & Nee, S. (2001). Phylogenetics and speciation. *TREE* **16**: 391-399.

Baverstock, P. R. & Moritz, C. (1996). Project design. In: *Molecular systematics.* (eds) D. M. Hillis, C. Moritz & B. K. Mable. Sinauer, MA.

Benton, M. J. (1999). Early origins of modern mammals: Molecules vs. morphology. *BioEssays* **21**: 1043-1051.

Birks, S. M. & Edwards, S. V. (2002). A phylogeny of the megapodes (Aves: Megapodiidae) based on nuclear and mitochondrial DNA sequences. *Mol. Phylogenet. Evol.* **23**: 408-421.

Birungi, J. & Arctander, P. (2001). Molecular systematics and phylogeny of the Reduncini (Artiodactyla: Bovidae) inferred from the analysis of mitochondrial cytochrome *b* gene sequences. *J. Mamm. Evol.* **8**: 125-147.

Bremer, K. (1988). The limits of amino acid sequence data in angiosperm phylogenetic reconstruction. *Evolution* **42**: 795-803.

Bremer, K. (1994). Branch support and tree stability. *Cladistics* **10**: 295-304.

Bromham, L., Penny, D., Rambaut, A. & Hendy, M. D. (2000). The power of relative rates tests depends on the data. *J. Mol. Evol.* **50**: 296-301.

Bronner, G. N., Hoffmann, M., Taylor, P. J., Chimimba, C. T., Best, P. B., Matthee, C. A. & Robinson, T. J. (2003). A revised systematic checklist of the extant mammals of the southern African subregion. *Ann. Nat. Mus.*

Broughton, R. E., Stanley, S. E. & Durrett, R. E. (2000). Quantification of homoplasy for nucleotide transitions and tranversions and a reexamination of assumptions in weighted phylogenetic analysis. *Syst. Biol.* **49**: 617-627.

Brown, W. M., Prager, E. M., Wang, A. & Wilson, A. C. (1982). Mitochondrial DNA sequences of primates: Tempo and mode of evolution. *J. Mol. Evol.* **18**: 225-239.

Buckland, R. A. & Evans, H. J. (1978). Cytogenetic aspects of phylogeny in Bovidae. I. G-banding. *Cytogenet. Cell. Genet.* **21**: 42-63.

Buckley, T. R., Simon, C. & Chambers, G. K. (2001). Exploring among-site variation models in a maximum likelihood framework using empirical data: The effects of model assumptions on estimates of topology, edge lengths, and bootstrap support. *Syst. Biol.* **50**: 67-86.

Buckley, T. R., Arensburger, P., Simon, C. & Chambers, G. K. (2002). Combining data, bayesian phylogenetics, and the origin of the New Zealand Circada genera. *Syst. Biol.* **51**: 4-18.

Bull, J. J, Huelsenbeck, J. P., Cunningham, C. W., Swofford, D. L. & Waddell, P. J. (1993). Partitioning and combining data in phylogenetic analysis. *Syst. Biol.* **42**: 384-397.

Cerling, T. E., Harris, J. M., MacFadden, B. J., Leakey, M. G., Quade, J., Eisenmann, V. & Ehleringer, J. R. (1997). Global vegetation change through the Miocene/Pliocene boundary. *Nature* **389**: 153-158.

Chevret, P., Denys, C., Jaeger, J. J., Michaux, J. & Catzeflis, F. M. (1993). Molecular evidence that the spiny mouse (Acomys) is more closely related to gerbils (Gerbillinae) than to true mice (Murinae). *Proc. Natl. Acad. Sci. USA.* **90**: 3433-3436.

Cortés-Ortiz, I., Bermingham, E., Rico, C., Rodriques-Luna, E., Sampaic, I. & Ruiz-Garcia, M. (2003). Molecular systematics and biogeography of the neotropical monkey genus, *Alouatta*. *Mol. Phylogenet. Evol.* **26**: 64-81.

Cunningham, C. W., Zhu, H. & Hillis, D. M. (1998). Best-fit maximum-likelihood models for phylogenetic inference: Empirical tests with known phylogenies. *Evolution* **52**: 978-987.

DeBry, R. W. (1992). The consistency of several phylogeny inference methods under varying evolutionary rates. *Mol. Biol. Evol.* **9**: 537-551.

Eldredge, N. & Gould, S. J. (1972). Punctuated equilibrium: An alternative to phyletic gradualism. In: *Models in Paleobioloigy*. (ed) T. J. M. Schopf. Freeman, San Francisco.

Ellerman, J. R., Morrison-Scott, T. C. S. & Hayman R. W. (1953). In: *Southern African mammals 1758-1951: A reclassification.* Trustees of the British Museum of Natural History, London.

Eriksson, T. (1998). AutoDecay. Version 4.0. Distributed by the author, Department of Botany, University of Stockholm, Sweden. Computor program available at http://www.bergianska.se/personal/TorstenE/.

Eriksson, T. (2001). AutoCladeS, version 0.2. Distributed by the author, Department of Botany, University of Stockholm, Sweden. Computor program available at http://www.bergianska.se/personal/TorstenE/.

Essop, M. F., Harley, E. H. & Baumgarten, I. (1997). A molecular phylogeny of some Bovidae based on restriction-site mapping of mitochondrial DNA. *J. Mam.* **78**: 377-386.

Estes, R. D. (1991). The significance of horns and other male secondary sexual characters in female bovids. *Appl. Anim. Behav. Sci.* **29**: 403-451.

Farris, J. S. (1989). The retention index and the rescaled consistency index. *Cladistics* **5**: 417-419.

Felsenstein, J. (1985). Confidence limits on phylogenies: An approach using the bootstrap evolution. *Genetics* **39**: 783-791.

Flagstad, Ø., Syvertsen, P. O., Stenseth, N. C. & Jakobsen, K. S. (2000). Environmental change and rates of evolution: The phylogeographic pattern within the hartebeest complex as related to climatic variation. *Proc. R. Soc. Lond. B* **268**: 667-677.

Flynn, J. J. & Nedbal, M. A. (1998). Phylogeny of the Carnivora (Mammalia): Congruence vs incompatibility among multiple data sets. *Mol. Phylogenet. Evol.* **9**: 414-426.

Fu, J. (2000). Toward the phylogeny of the family Lacertidae - Why 4708 base pairs of mitochondrial DNA sequences cannot draw the picture. *Biol. J. Linn. Soc.* **71**: 203-217.

Gagnon, M. & Chew, A. E. (2000). Dietary preferences in extant African Bovidae. *J. Mam.* **8**: 490-511.

Gallagher, D. S. & Womack, J. E. (1992). Chromosome conservation in the Bovidae. *J. Hered.* **83**: 287-298.

Gatesy, J., Yelon, D., DeSalle, R. & Vrba, E. S. (1992). Phylogeny of the Bovidae (Artiodactyla, Mammalia), based on mitochondrial ribosomal DNA sequences. *Mol. Biol. Evol.* **9**: 433-446.

Gatesy, J., Hayashi, C., Cronin, M. A. & Arctander, P. (1996). Evidence from milk casein genes that cetaceans are close relatives of hippopotamid artiodactyls. *Mol. Biol. Evol.* **13**: 954-963.

Gatesy, J., Amato, G., Vrba, E. S., Schaller, G. & DeSalle, R. (1997). A cladistic analysis of mitochondrial ribosomal DNA from the Bovidae. *Mol. Phylogenet. Evol.* **7**: 303-319.

Gatesy, J., Milinkovitch, M., Waddell, V. & Stanhope, M. (1999). Stability of cladistic relationships between Cetacea and higher level artiodactyl taxa. *Syst. Biol.* **48**: 6-20.

Gatesy, J. & Arctander, P. (2000). Hidden morphological support for the phylogenetic placement of *pseudoryx nghetinhensis* with bovine bovids: A combined analysis of gross anatomical evidence and DNA sequences from five genes. *Syst. Biol.* **3**: 515-538.

Gatesy, J., Matthee, C. A., DeSalle, R. & Hayashi, C. (2002). Resolution of a supertree/supermatrix paradox. *Syst. Biol.* **51**: 652-664.

Gentry, A. W. (1978). In: *Evolution of African mammals*. Harvard University Press, Cambridge.

Gentry, A. W. (1990). Evolution and the dispersal of African Bovidae. In: *Horns, proghorns and antlers: Evolution, morphology, physiology and social significance.* (eds) G. A. Bubenik & A. B. Bubenik. Springer- Verlag, New York.

Gentry, A. W. (1992). The subfamilies and tribes of the family Bovidae. *Mammal Rev.* **22**: 1-32.

Georgiadis, N. J., Kat, P. W., Oketch, H. & Patton, J. (1990). Allozyme divergence within the Bovidae. *Evolution* **44**: 2135-2149.

Goldman, N., Anderson, J. P. & Rodrigo, A. G. (2000). Likelihood-based tests of topologies in phylogenetics. *Syst. Biol.* **49**: 652-670.

Goloboff, P. A. (1993). Estimating character weights during tree search. *Cladistics* **9**: 83-91.

Graur, D. & Higgins, D. (1994). Molecular evidence for the inclusion of cetaceans within the order Artiodactyla. *Mol. Biol. Evol.* **11**: 357-364.

Gray, J. E. (1821). On the natural arrangement of vertebrate mammals. *Lond. Med. Repos.* **15**: 296-310.

Halanych, K. M., Dembaski, J. K., Jansen van Vuuren, B., Klein, D. R. & Cook, J. A. (1999). Cytochrome *b* phylogeny of North American hares and jackrabbits (*Lepus*, Logomorpha) and the effects of saturation in ingroup taxa. *Mol. Phylogenet. Evol.* **11**: 213-221.

Haltenorth, T. (1963). Klassifikation der säugetiere: Artiodactyla I. In: *Handbuch der zoologie* **8**: 1-167.

Harrison, R. G. (1991). Molecular changes at speciation. *Ann. Rev. Ecol. Syst.* **22**: 281-308.

Hasegawa, M. (1990). Mitochondrial DNA evolution in primates: Transition rate has been extremely low in the lemur. *J. Mol. Evol.* **31**: 113-121.

Hassanin, A. & Douzery, E. J. P. (1999a). The tribal radiation of the family Bovidae (Artiodactyla) and the evolution of the mitochondrial cytochrome *b* gene. *Mol. Phylogenet. Evol.* **13**: 227-243.

Hassanin, A. & Douzery, E. J. P. (1999b). Evolutionary affinities of the enigmatic saola (*Pseudoryx nghetinhensis*) in the context of the molecular phylogeny of Bovidae. *Proc. R. Soc. Lond.* **266**: 893-900.

Hassanin, A. & Douzery, E. J. P. (2003). Molecular and morphological phylogenies of ruminantia and the alternative position of Moschidae. *Syst. Biol.* **52**: 206-228.

Hill, A., Drake, R., Tauxe, L., Monaghan, M., Barry, J. C., Behrensma, A. K., Curtis, G., Jacobs, B. F., Jacobs, L. & Johnson, N. (1985). Neogene paleontology and geochronology of Baringo Basin, Kenya. *J. Hum. Evol.* **14**: 759-773.

Hillis, D. M. (1987). Molecular vs morphological approaches to systematics. *Ann. Rev. Ecol. Syst.* **18**: 23-42.

Hillis, D. M. & Dixon, M. T. (1991). Ribosomal DNA: Molecular evolution and phylogenetic inference. *Quart. Rev. Biol.* **66**: 411-453.

Hillis, D. M. & Bull, J. J. (1993). An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. *Syst. Biol.* **42**: 182-192.

Hillis, D. M., Mable, B. K. & Moritz, C. (1996). Applications of molecular systematics: The state of the field and a look to the future. In: *Molecular Systematics*. (eds) D. M. Hillis, C. Moritz & B. K. Mable. pp. 515-543. Sinauer, Sunderland, MA.

Hillis, D. M. (1999). SINEs of the perfect character. *Proc. Natl. Acad. Sci. USA* **96**: 9979-9981.

Huelsenbeck, J. P. & Hillis, D. (1993). Success of phylogenetic methods in the four-taxon case. *Syst. Biol.* **42**: 247-264.

Huelsenbeck, J. P. (1995). Performance of phylogenetic methods in sumulations. *Syst. Biol.* **44**: 17-48.

Huelsenbeck, J. P. & Ronquest, F. (2001). MrBayes: Bayesian inference of phylogeny. *Bioinformatics* **17**: 754-755.

Irwin, D. M., Kocher, T. D. & Wilson, A. C. (1991). Evolution of cytochrome *b* gene of mammals. *J. Mol. Evol.* **32**: 128-144.

Janeck, L. L., Honeycutt, R. L., Adkins, R. M. & Davis, S. K. (1996). Mitochondrial gene sequences and the molecular systematics of the Artiodactyla subfamily Bovinae. *Mol. Phylogenet. Evol.* **6**: 107-119.

Janis, C. M. & Scott, K. M. (1987). The interrelationships of higher ruminant families with special emphasis on the members of the Cervoidea. *Am. Mus. Novit.* **2893**: 1-85.

Jones, K. E., Purvis, A., MacLarnon, A., Bininda-Emonds, O. R. P. & Simmons, N. B. (2002). A phylogenetic supertree of the bats (Mammalia: Chiroptera). *Biol. Rev.* **77**: 223-259.

Källersjö, M., Albert, V. A. & Farris, J. S. (1999). Homoplasy *increases* phylogenetic structure. *Cladistics* **15**: 91-93.

Kelsey, C. R., Crandall, K. A. & Voevodin, A. F. (1999). Different models, different trees: The geographic origin of PTLV-I. *Mol. Phylogenet. Evol.* **13**: 336-347.

Kingdon, J. (1989). Bovids, horned ungulates. In: *East African mammals: An atlas of evolution in Africa, vol. IIIC.* Academic Press, London.

Kingdon, J. (1997). In: *The Kingdon field guide to African mammals.* Academic Press, London.

Kluge, A. G. & Farris, J. S. (1969). Quantitative phylogenetics and the evolution of Anurans. *Syst. Zool.* **18**: 1-32.

Kluge, A. G. (1989). A concern for evidence and a phylogenetic hypothesis of relationships among *Epicrates* (Boidae, Serpentes). *Syst. Zool.* **38**: 7-25.

Kocher, T. D., Thomas, W. K., Meyer, A., Edwards, S. V., Pääbo, S., Villablanca, F. X. & Wilson, A. C. (1989). Dynamics of mitochondrial DNA evolution in animals: Amplification with conserved primers. *Proc. Natl. Acad. Sci. USA.* **86**: 6196-6200.

Kumar, S. & Hedges, S. B. (1998). A molecular timescale for vertebrate evolution. *Nature* **392**: 917-920.

Kuhner, M. K. & Felsenstein, J. (1994). A simulation comparison of phylogeny algorithms under equal and unequal evolutionary rates. *Mol. Biol. Evol.* **11**: 459-468.

Lara, M. C., Patton, J. L. & DaSilva, M. N. F. (1996). The similtaneous diversification of South American echimyid rodents (*Hystricognathi*) based on complete cytochrome *b* sequences. *Mol. Phylogenet. Evol.* **5**: 403-413.

Leaché, A. D. & Reeder, T. W. (2002). Molecular systematics of the eastern fence lizard (*Sceloporus undulates*): A comparison of parsimony, likelihood, and bayesian approaches. *Syst. Biol.* **51**: 44-68.

Lee, M. S. Y. (2000). Tree robustness and clade significance. *Syst. Biol.* **49**: 829-836.

Lee, M. S. Y. (2001). Uninformative characters and apparent conflict between molecules and morphology. *Mol. Biol. Evol.* **18**: 676-680.

Lee, M. S. Y. & Hugall, A. F. (2003). Partitioned likelihood support and the evaluation of data set conflict. *Syst. Biol.* **52**: 15-22.

Lindsay, J. A. (1998). Past climates of southern Africa. In: *Climates of the southern continents: Present, past and future*. (eds) H. A. Bridgeman, J. E. Hobbs & J. A. Lindsay. John Wiley and sons Ltd, England.

Lowenstein, J. M. (1986). Bovid relationships based on serum immunology. *S. Afr. J. Sci.* **82**: 77-78.

Lydekker, R. (1910). The spotted kudu. *Nature* **84**: 397-398.

Lydekker, R. (1911). On the mountain nyala, T*ragelaphus buxtoni. Proc. R. Soc. Lond.* **21**: 348-353.

Lyons-Weiler, J. & Milinkovitch, M. C. (1997). A phylogenetic approach to the problem of differential lineage sorting. *Mol. Biol. Evol.* **14**: 968-975.

Macnish, M. G., Morgan-Ryan, U. M., Monis, P. T., Behnke, J. M. & Thompson, R. C. (2002). A molecular phylogeny of nuclear and mitochondrial sequences in *Hymenolepis nana* (Cestoda) supports the existence of a cryptic species. *Parasitology* **125**: 567-575.

Maddison, W. P. & Maddison, D. R. (1992). MacClade: Analysis of phylogeny and character evolution, version4.0. Sinauer, Sunderland, MA.

Madsen, O., Scally, M., Douady, C. J., Kao, D. J., DeBry, R. W., Adkins, R., Amrine, H. M., Stanhope, M. J., de Jong, W. W. & Springer, M. S. (2001). Parallel adaptive radiations in two major clades of placental mammals. *Nature* **409**: 610-614.

Manceau, V., Després, L., Bouvet, J. & Taberlet, P. (1999). Systematics of the genus *Capra* inferred from mitochondrial DNA sequence data. *Mol. Phylogenet. Evol.* **13**: 504-510.

Matthee, C. A. & Robinson, T. J. (1997). Molecular phylogeny of the springhare, *Pedetes capensis*, based on mitochondrial DNA sequences. *Mol. Biol. Evol.* **14**: 20-29.

Matthee, C. A. & Robinson, T. J. (1999a). Cytochrome *b* phylogeny of family Bovidae: Resolution within the Alcephini, Antilopini, Neotragini and Tragelaphini. *Mol. Phylogenet. Evol.* **12**: 31-46.

Matthee, C. A. & Robinson, T. J. (1999b). Mitochondrial DNA population structure of roan and sable antelope: Implications for the translocation and conservation of species. *Mol. Ecol.* **8**: 227-238.

Matthee, C. A., Burzlaft, J. D., Taylor, J. F. & Davis, S. K. (2001). Mining the mammalian genome for artiodactyl systematics. *Syst. Biol.* **50**: 367-390.

Matthee, C. A. & Davis, S. K. (2001). Molecular insights into the evolution of the family Bovidae: A nuclear DNA perspective. *Mol. Biol. Evol.* **18**: 1220-1230.

Mau, B. (1996). Bayesian phylogenetic inference via Markov Chain Monte Carlo methods. Ph. D. Dissertation, Univ. of Wisconsin, Madison.

Mau, B. & Newton, M. (1997). Phylogenetic inference for binary data on dendrograms using Markov Chain Monte Carlo. *J. Compt. Graph. Stat.* **6**: 122-131.

Mau, B., Newton, M. & Larget, B. (1999). Bayesian phylogenetic inference via Markov Chain Monte Carlo methods. *Biometrics* **55**: 1-12.

Milinkovitch, M. C., Leduc, R. G., Adachi, J., Farnir, F., Georges, M. & Hasegawa, M. (1996). Effects of character weighting and species sampling on phylogeny reconstruction: A case study based on cetaceans. *Genetics* **144**: 1817-1833.

Milinkovitch, M. C. & Lyons-Weiler, J. (1998). Finding optimal ingroup topologies and convexities when the choice of outgroup is not obvious. *Mol. Phylogenet. Evol.* **9**: 348-357.

Miyamoto, M. M. & Fitch, W. M. (1995). Testing species phylogenies and phylogenetic methods with congruence. *Syst. Biol.* **44**: 64-76.

Modi, W. S., Gallagher, D. S. & Womack, J. E. (1996). Evolutionary histories of highly repeated DNA families among the artiodactyla (Mammalia). *J. Mol. Evol.* **42**: 337-349.

Montgelard, C., Catzeflis, F. M. & Douzery, E. (1997). Phylogenetic relationships of artiodactyls and cetaceans as deduced from the comparison of cytochrome *b* and 12S rRNA mitochondrial sequences. *Mol. Biol. Evol.* **14**: 550-559.

Moritz, C. & Hillis, D. M. (1996). Molecular systematics: Context and controversies. In: *Molecular systematics.* (eds) D. M. Hillis, C. Moritz, & B. K. Mable. Sinauer, MA.

Morozov, P., Sitnikova, T., Churchill, G., Ayala, F. J. & Rzhetsky, A. (2000). A new method for characterizing replacement rate variation in molecular sequences: Application of the Fourier and Wavelet models to drosophilia and mammalian proteins. *Genetics* **154**: 381-395.

Mullis, K. B. & Faloona, F. A. (1987). Specific synthesis of DNA in vitro *via* a polymerase catalyzed chain reaction. *Meth. Enzymol.* **155**: 335-350.

Murphy, W. J., Eizirik, E., Johnson, W. E., Zhang, Y. P., Ryder, O. A. & O'Brien, S. J. (2001). Molecular phylogenetics and the origins of placental mammals. *Nature* **409**: 614-618.

Nei, M. & Kumar, S. (2000). In: *Molecular evolution and phylogenetics.* Oxford University Press, New York.

Nixon, K. C. & Carpenter, J. M. (1993). On outgroups. *Cladistics* **9**: 413-426.

Nowak, R. M. (1999). In: *Walker's Mammals of the World. Edn XI.* The Johns Hopkins University Press, Baltimore and London.

O'Grady, P. M. & Kidwell, M. G. (2002). Phylogeny of the subgenus *Sophophora* (Diptera: Drosophilidae) based on combined analysis of nuclear and mitochondrial sequences. *Mol. Phylogenet. Evol.* **22**: 442-453.

Pamilo, P. & Nei, M. (1988). The relationships between gene trees and species trees. *Mol. Biol. Evol.* **5**: 568-583.

Pasada, D. & Crandall, K. A. (1998). MODELTEST version 3.06. Brigham Young University.

Pasada, D. & Crandall, K. A. (2001). Selecting the best-fit model of nucleotide substitution. *Syst. Biol.* **50**: 580-601.

Pitra, C., Fürbass, R. & Seyfert, H. M. (1997). Molecular phylogeny of the tribe Bovini (Mammalia: Artiodactyla): Alternative placement of the Anoa. *J. Evol. Biol.* **10**: 589-600.

Rannala, B. & Yang, Z. (1996). Probability distribution of molecular evolutionary trees: A method of phylogenetic inference. *J. Mol. Evol.* **43**: 304-311.

Rebholz, W. & Harley, E. (1999). Phylogenetic relationships in the bovid subfamily Antilopinae based on mitochondrial DNA sequences. *Mol. Phylogenet. Evol.* **12**: 87-94.

Remigio, E. A. & Hebert, P. D. N. (2000). Affinities among anostracan (Crustacea: Branchiopoda) families inferred from phylogenetic analysis of multiple gene sequences. *Mol. Phylogenet. Evol.* **17**: 117-128.

Rice, W. R. (1989). Analyzing tables of statistical tests. *Evolution* **43**: 223-225.

Roberts, A. (1952). In: *The mammals of South Africa.* Trustees of "The Mammals of South Africa" book fund, Johannesburg.

Robinson, M., Gouy, M., Gautier, C. & Mouchirod, D. (1998). Sensitivity of relative rates tests to taxonomic sampling. *Mol. Biol. Evol.* **15**: 1091-1098.

Robinson, T. J., Harrison, W. R., Ponce de León, F. A., Davis, S. K. & Elder, F. F. B. (1998). A molecular cytogenetic analysis of X chromosome repatterning in the Bovidae: Transpositions, inversions, and phylogenetic inference. *Cytogenet. Cell. Genet.* **80**: 179-184.

Robinson-Rechavi, M. & Huchon, P. (2000). RRTree: Relative-rates test between groups of sequences on a phylogenetic tree. *Bioinformatics* **16**: 296-297.

Rodríguez, F. J., Oliver, J. L., Marín, A. & Medina, J. R. (1990). The general stochastic model of nucleotide substitution. *J. Theor. Biol.* **142**: 485-501.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). In*: Molecular cloning: A laboratory manual, Edn II.* Cold Spring Harbour Laboratory Press, New York.

Sanderson, M. J. & Donoghue, M. J. (1989). Patterns of variation in levels of homoplasy. *Evolution* **43**: 1781-1795.

Sanderson, M. J. (1995). Objections to bootstrapping phylogenies: A critique. *Syst. Biol.* **44**: 299-320.

Sanderson, M. J. & Donoghue, M. J. (1996). The relationship between homoplasy and confidence in a phylogenetic tree. In: *Homoplasy: The recurrence of similarity in evolution.* (eds) M. J. Sanderson & L. Hufford. Academic Press, San Diego.

Schwabland, J. (1993). In: *Safari club international record book of trophy animals, Africa field edn, Edn VII.* Safari Club International/Walsworth Publishing Company.

Schwarz, G. (1974). Estimating the dimension of a model. *Ann. Stat.* **6**: 461-464.

Shimodaira, H. & Hasegawa, M. (1999). Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Mol. Biol. Evol.* **16**: 1114-1116.

Simon, D. & Larget, B. (2000). Bayesian analysis in molecular biology and evolution (BAMBE), version 2.03 beta. Department of mathematics and computer science, Duquesne University.

Simpson, G. G. (1945). The principles of classification and a classification of mammals. *Bull. Am. Mus, Nat. Hist.* **85**: 1-350.

Slade, R. W., Moritz, C. & Heideman, A. (1994). Multiple nuclear-gene phylogenies: Application to pinnipeds and comparison with a mitochondrial DNA gene phylogeny. *Mol. Biol. Evol.* **11**: 341-356.

Smith, A. B. (1994). Rooting molecular trees: Problems and strategies. *Biol. J. Linn. Soc.* **51**: 279-292.

Springer, M. S., DeBry, R. W., Douady, C., Amrine, H. M., Madsen, O., de Jong, W. W. & Stanhope, M. J. (2001). Mitochondrial versus nuclear gene sequences in deep-level mammalian phylogeny reconstruction. *Mol. Biol. Evol.* **18**: 132-143.

Sullivan, J. (1996). Combining data with different distributions of among-site rate variation. *Syst. Biol.* **45**: 375-380.

Sullivan, J. & Swofford, D. L. (1997). Are guinea pigs rodents? The importance of adequate models in molecular phylogenies. *J. Mammal. Evol.* **4**: 77-86.

Suzuki, Y., Glazko, G. V. & Nei, M. (2002). Overcredibility of molecular phylogenies obtained by bayesian phylogenetics. *PNAS* **99**: 16138-16143.

Swofford, D. L., Olsen, G. J., Waddell, P. J. & Hillis, D. M. (1996). Phylogenetic inference. In: *Molecular systematics.* (eds) D. M. Hillis, C. Moritz, & B. K. Mable. Sinauer, MA.

Swofford, D. L. (1999). PAUP: Phylogenetic Analysis using Parsimony. Version 4.0. Illinois Natl. Hist. Surr, Champaign.

Tajima, F. (1993). Simple methods for testing the molecular evolutionary clock hypothesis. *Genetics* **135**: 599-607.

Takezaki, N., Rzhetsky, A. & Nei, M. (1995). Phylogenetic test of the molecular clock and linearized trees. *Mol. Biol. Evol.* **12**: 823-833.

Tamura, K. (1994). Model selection in the estimation of the number of nucleotide substitutions. *Mol. Biol. Evol.* **11**: 154-157.

Templeton, A. R. (1983). Phylogenetic inference from restriction site endonuclease cleavage site maps with particular reference to the humans and apes. *Evolution* **37**: 221-244.

Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. (1997). The Clustal X windows interface: Flexible strategies for multiples sequence alignment aided by quality analysis tools. *Nucl. Acids Res.* **25**: 4876-4882.

Tijskens, J. (1968). Preliminary notes on the F1 bongo antelope X sitatunga hybrids, *Taurotragus eurycerus* X *Tragelaphus spekei* at Antwerp zoo. *Int. zoo. Yb.* **8**: 137-139.

Vrba, E. S. (1979). Phylogenetic analysis and classification of fossil and recent Alcelaphini (Mammalia: Bovidae). *Biol. J. Linn. Soc.* **11**: 207-228.

Vrba, E. S. (1980). Evolution, species and fossils: How does life evolve? *S. Afr. J. Sci.* **76**: 61-83.

Vrba, E. S. (1984). What is species selection? *Syst. Zool.* **33**: 318-328.

Vrba, E. S. (1985a). African Bovidae: Evolutionary events since the Miocene. *S. Afr. J. Sci.* **81**: 263-266.

Vrba, E. S. (1985b). Environment and evolution: Alternative causes of the temporal distribution of evolutionary events. *S. Afri. J. Sci.* **81**: 229-236.

Vrba, E. S. (1987). Ecology in relation to speciation rates. Some case histories of Miocene-recent mammal clades. *Evol. Ecol.* **1**: 283-300.

Vrba, E. S. (1995). The fossil record of African antelopes (Mammalia, Bovidae) in relation to human evolution and paleoclimate. In: *Paleoclimate and evolution, with emphasis on human origins* (eds) E. S. Vrba, G. H. S. Denton, T. C. Partridge & L. H. Burckle. pp. 385-424. New Haven, CT and London: Yale University Press.

Waddell, P. J., Kishino, H. & Ota, R. (2001). A phylogenetic foundation for comparative mammalian genomics. *Genome Informatics* **12**: 141-154.

Walker, E. P (1964). In: *Mammals of the World*. pp. 1412-1420. Johns Hopkins Press, Baltimore.

Wilson, P. E. & Reeder, D. M. (1993). In: *Mammal species of the world: A taxonomic and geographic reference*. Smithsonian Institute Press, London.

Yang, Z. (1994). Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: Appropriate methods. *J. Mol. Evol.* **39**: 306-314.

Yang, Z. & Rannala, B. (1997). Bayesian phylogenetic inference using DNA sequences: A Markov Chain Monte Carlo method. *Mol. Biol. Evol.* **14**: 717-724.

Yoder, A. D., Irwin, J. A. & Payseur, B. A. (2001). Failure of the ILD to determine data combinability for slow loris phylogeny. *Syst. Biol.* **50**: 408-424.

Zachos, J., Pagani, M., Sloan, L., Thomas, E. & Billups, K. (2001). Trends, rhythms, and aberrations in global climate 65 Ma to present. *Science* **292**: 686-693.

Zuckerkandl, E. & Pauling, L. (1965). Evolutionary divergence and convergence in proteins. In: *Evolving genes and proteins*. (eds) V. Bryson & H. J. Vogel. New York: Academic.

Zwickl, D. J. & Hillis, D. M. (2002). Increased taxon sampling greatly reduces phylogenetic error. *Syst. Biol.* **51**: 588-598.

Appendix: Alignment of the sequence data used in the present study, (-) representing a gap and (?) representing missing data. Character partitions are as follows: 1-1140 cytochrome b, 1141-1923 cox III, 1924-2117 cyp 19, 2118-2458 lactoferrin, 2459-2806 16S rRNA, 2807-3397 12S rRNA, 33398-3894 PRKCI, 3895-4576 MGF, 4577-5243 THY, 5244-6008 SPTBN. Taxonomic sampling is as follows: GE- *T. derbianus*, CE- *T. oryx*, GK- *T. strepsiceros*, LK- *T. imberbis*, CN- *T. angasi*, MN- *T. buxtoni*, B- *T. euryceros*, SIT- *T. spekei*, BB- T. scriptus, IMP- *A. melapus*, COW- *B. taurus*, BUF- *S. caffer*, NIL- *B. tragocamelus.*

```
          10        20        30        40        50        60        70        80
          .         .         .         .         .         .         .         .
GE1       ATGACCAACATTCGAAAATCCCACCCACTAATAAAAATTGTAAACAACGCATTCATCGACCTCCCAGCCCCATCGAACAT    80
CE1       ATGACCAACATTCGAAAATCCCACCCACTAATAAAAATTGTAAACAACGCATTCATCGACCTCCCAACCCCATCAAACAT    80
CE2       ATGACCAACATTCGAAAATCCCACCCACTAATAAAAATTGTAAACAACGCATTCATCGACCTCCCAACCCCATCAAACAT    80
GK1       ATGACCAACATTCGAAAATCCCACCCACTAATAAAAATTGTAAACAACGCATTCATCGACCTCCCAGCCCCATCAAATAT    80
GK2       ATGACCAACATTCGAAAATCCCACCCACTAATAAAAATTGTAAACAACGCATTCATCGACCTCCCAGCCCCATCAAATAT    80
LK1       ATGATCAACATTCGAAAATCTCACCCACTAATAAAAATTGTAAACAACGCATTCATTGACCTCCCAACCCCACCAAACAT    80
CN1       ATGACCAACATCCGAAAATCCCACCCACTGATAAAAATTGTAAACAACGCATTCATCGACCTCCCAGCCCCATCAAACAT    80
CN2       ATGACCAACATCCGAAAATCCCACCCACTGATAAAAATTGTAAACAACGCATTCATCGACCTCCCAGCCCCATCAAACAT    80
MN1       ATGACCAACATTCGAAAATCCCACCCACTAATAAAAATTGTAAATAACGCATTCATTGACCTCCCAGCCCCATCAAACAT    80
MN2       ATGACCAACATTCGAAAATCCCACCCACTAATAAAAATTGTAAATAACGCATTCATTGACCTCCCAGCCCCATCAAACAT    80
B1        ATGATCAACATTCGAAAATCCCACCCACTGATAAAAATTGTAAACAACGCATTCATTGACCTCCCAGCCCCATCAAACAT    80
B2        ATGATCAACATTCGAAAATCCCACCCACTGATAAAAATTGTAAACAACGCATTCATTGACCTCCCAGCCCCATCAAACAT    80
SIT1      ATGACCAACATTCGAAAATCCCACCCACTGATAAAAATTGTAAACAACGCATTTATTGACCTCCCAGCCCCATCAAATAT    80
BB1       ATGACCAACATCCGAAAATCCCACCCACTGATAAAAATTGTAAACAACGCATTCATCGACCTCCCAGCCCCATCAAACAT    80
BB2       ATGACCAACATCCGAAAATCCCACCCACTGATAAAAATTGTAAACAACGCATTCATCGACCTCCCAGCCCCATCAAACAT    80
IMP       ATGACCACGATCTGAAAAACCCACCCACTAATAAAAATTGATAATAACGCATTCATTGACCTACCAGCCCCATCAAACAT    80
COW       ATGACTAACATTCGAAAGTCCCACCCACTAATAAAAATTGTAAACAATGCATTCATCGACCTTCCAGCCCCATCGAACAT    80
BUF       ATGACCAATATTCGAAAATCCCACCCACTAATAAAAATTCTAAACAACGCCTTCATTGATCTCCCAGCTCCATCAAACAT    80
NIL       ATGACTAACATCCGAAAATCCCACCCATTAATAAAAATTGTAAACAATGCATTCATTGACCTCCCAGCCCCATCGAATAT    80

          90        100       110       120       130       140       150       160
          .         .         .         .         .         .         .         .
GE1       CTCATCATGATGAAACTTCGGATCCCTTCTAGGAATCTGCCTGATTCTGCAAATCCTCACAGGCTTATTTCTAGCCATAC    160
CE1       CTCATCATGATGAAACTTCGGATCCCTCCTAGGAATCTGCCTAATTCTACAAATCCTCACAGGCTTATTCCTAGCTATAC    160
CE2       CTCATCATGATGAAACTTCGGATCCCTCCTAGGAATCTGCCTAATTCTACRAATCCTCACGGGCTTATTCCTAGCTATAC    160
GK1       CTCATCATGATGAAACTTCGGATCCCTTCTAGGAATCTGCCTAATTCTACAAATCCTCACAGGCTTATTCCTAGCAATAC    160
GK2       CTCATCATGATGAAACTTCGGATCCCTTCTAGGAATCTGCCTAATTCTACAAATCCTCACAGGCTTATTCCTAGCAATAC    160
LK1       CTCATCATGATGAAATTTCGGATCACTTCTAGGGATCTGCCTAGTTCTGCAAATCCTTACAGGTTTATTTCTAGCAATAC    160
CN1       CTCTTCATGATGAAACTTCGGATCCCTCTTAGGAGTCTGCCTAATTCTACAAATCTTCACAGGCTTATTCCTGGCCATAC    160
CN2       CTCTTCATGATGAAACTTCGGATCCCTCTTAGGAGTCTGCCTAATTCTACAAATCCTTACAGGCTTATTCCTGGCCATAC    160
MN1       CTCATCATGATGAAACTTCGGATCCCTTCTAGGAGTCTGCCTAATCCTGCAAATCCTCACAGGCCTATTCCTAGCTATAC    160
MN2       CTCATCATGATGAAACTTCGGATCCCTTCTAGGAGTCTGCCTAATCCTGCAAATCCTCACAGGCCTATTCCTAGCTATAC    160
B1        CTCATCATGATGAAACTTCGGATCCCTTCTAGGAATCTGCCTAATTCTACAAATCCTCACAGGCTTATTCCTAGCTATAC    160
B2        CTCATCATGATGAAACTTCGGATCCCTTCTAGGAATCTGCCTAATTCTACAAATCCTCACAGGCTTATTCCTAGCTATAC    160
SIT1      CTCATCATGATGAAACTTCGGGTCCCTTCTAGGAATCTGCCTAATTCTCCAAATCCTCACAGGCTTATTTCTAGCTATAC    160
BB1       CTCTTCATGATGAAACTTCGGGTCCCTCTTAGGGATCTGCCTAATTCTGCAAATCCTCACGGGCTTATTTCTAGCTATAC    160
BB2       CTCATCATGATGAAACTTCGGGTCCCTCCTAGGGATCTGCCTAATTCTGCAAATCCTCACAGGCTTATTTCTAGCTATAC    160
IMP       TTCATCATGATGAAACTTCGGTTCTCTCCTAGGCATCTGCCTAATCCTACAAATCCTAACAGGCCTATTTCTAGCAATAC    160
COW       TTCATCATGATGGAATTTCGGTTCCCTCCTGGGAATCTGCCTAATCCTACAAATCCTCACAGGCCTATTCCTAGCAATAC    160
BUF       CTCATCATGATGAAATTTTGGCTCCCTCCTAGGTATTTGCCTAATCTTACAAATCCTCACAGGCTTATTCCTAGCAATAC    160
NIL       CTCATCATGATGAAACTTTGGCTCCCTACTAGGGATCTGCCTAATCCTACAAATTCTCACAGGCCTATTTTTAGCTATAC    160

          170       180       190       200       210       220       230       240
          .         .         .         .         .         .         .         .
GE1       ACTACACATCAGACACAACAACAGCATTCTCCTCTGTCACCCACATCTGTCGAGACGTAAACTATGGCTGAATCATCCGA    240
CE1       ACTACACATCAGACACAACAACAGCATTCTCCTCTGTCACTCACATCTGCCGAGACGTAAACTATGGCTGAATCATCCGA    240
CE2       ACTACACATCAGACACAACAACAGCATTCTCCTCTGTCACTCACATCTGCCGAGACGTAAACTATGGCTGAATCATCCGA    240
GK1       ACTACACGTCAGACACAACAACAGCATTCTCTTCTGTCACTCATATCTGCCGAGACGTAAACTACGGCTGAATTATCCGA    240
GK2       ACTACACGTCAGACACAACAACAGCATTCTCTTCTGTCACTCATATCTGCCGAGACGTAAACTACGGCTGAATTATCCGA    240
LK1       ACTACACATCAGACACAATAACAGCATTTTCCTCCGTCACCCACATCTGTCGAGACGTAAACTACGGCTGAATCATCCGG    240
CN1       ACTACACATCAGACACAATAACAGCATTTTCCTCCGTCACCCACATCTGTCGAGATGTAAACTACGGCTGAATTATCCGA    240
CN2       ACTACACATCAGACACCATAACAGCATTTTCCTCCGTCACCCACATCTGTCGAGATGTAAACTACGGCTGAATTATCCGA    240
MN1       ACTACACGTCAGACACAACAACAGCATTCTCCTCTGTCACCCATATCTGCCGAGACGTAAACTACGGCTGAATTATCCGA    240
MN2       ACTACACGTCAGACACAACAACAGCATTCTCCTCTGTCACCCATATCTGCCGAGACGTAAACTACGGCTGAATTATCCGA    240
B1        ACTACACGTCAGACACAACAACAGCATTCTCCTCTGTTACCCATATCTGCCGAGACGTAAATTACGGCTGAATCATTCGG    240
B2        ACTACACGTCAGACACAACAACAGCATTCTCCTCTGTTACCCATATCTGCCGAGACGTAAATTACGGCTGAATCATTCGA    240
SIT1      ACTACACGTCAGACACAACAACAGCATTCTCCTCTGTCACCCATATCTGCCGAGACGTAAACTACGGCTGAATTATCCGA    240
BB1       ACTACACATGAGACACAATAACAGCATTCTCCTCTGTCACCCATATCTGCCGGGACGTAAACTACGGCTGGATTATCCGA    240
BB2       ACTACACATCAGACACAATAACAGCATTCTCCTCTGTCACCCATATCTGCCGAGACGTAAACTACGGCTGAATTATCCGA    240
IMP       ACTATACATCTGACACAACAACAGCATTTTCCTCTGTCACCCATATTTGCCGAGATGTTAATTACGGATGGATTATCCGA    240
COW       ACTACACATCCGACACAACAACAGCATTCTCCTCTGTTACCCATATCTGCCGAGACGTGAACTACGGCTGAATCATCCGA    240
BUF       ATTATTCATCAGATACAACAACAGCATTCTCTTCCGTCGCCCACATCTGCCGAGACGTTAACTACGGATGAATTATTCGA    240
NIL       ACTACACATCCGACACAATAACAGCATTCGCTTCTGTCACCCACATTTGCCGAGACGTTAACTATGGCTGAATTATCCGA    240
```

```
                250       260       270       280       290       300       310       320
                .         .         .         .         .         .         .         .
GE1     TACATACATGCAAACGGAGCCTCAATATTTTTCATTTGCTTATATATACATGTAGGACGGGGGATATACTACGGGTCATA  320
CE1     TATATACATGCAAACGGAGCCTCAATATTTTTCATCTGCCTATACATACACGTGGGACGAGGAATATATTACGGATCATA  320
CE2     TATATACATGCAAACGGAGCCTCAATATTTTTCATCTGCCTATACATACACGTGGGACGAGGAATATATTACGGATCATA  320
GK1     TATATACATGCAAACGGAGCCTCAATATTCTTCATTTGCCTGTATGTACATGTAGGACGGGGAATATATTACGGATCATA  320
GK2     TATATACATGCAAACGGAGCCTCAATATTCTTCATTTGCCTGTATGTACATGTAGGACGGGGAATATATTACGGATCATA  320
LK1     TATATGCATGCAAATGGAGCTTCAATATTTTTCATTTGCTTATATATGCACGTAGGACGAGGACTGTACTACGGGTCATA  320
CN1     TACATACATGCAAACGGAGCCTCAATATTTTTTATCTGCCTATATATACACGTAGGACGAGGATTATATTATGGGTCTTA  320
CN2     TACATACATGCAAACGGAGCCTCAATATTTTTTATCTGCCTATATATACACGTAGGACGAGGATTATATTATGGGTCTTA  320
MN1     TATATACATGCAAATGGGGCCTCAATATTCTTCATTTGCCTGTACATACATGTAGGACGAGGAATATACTACGGGTCCTA  320
MN2     TATATACATGCAAATGGGGCCTCAATATTCTTCATTTGCCTGTACATACATGTAGGACGAGGAATATACTACGGGTCCTA  320
B1      TATATACATGCAAATGGAGCCTCAATATTCTTCATCTGCCTGTACATACATGTAGGACGAGGAATATACTACGGATCCTA  320
B2      TACATACATGCAAATGGAGCCTCAATATTCTTCATCTGCCTGTACATACATGTAGGACGAGGAATATACTACGGATCCTA  320
SIT1    TATATACATGCAAATGGAGCCTCAATATTCTTCATCTGCCTATATATACATGTAGGACGAGGAATATACTACGGATCCTA  320
BB1     TACATACACGCAAATGGAGCCTCAATATTCTTCATTTGCCTGTATATGCATGTAGGACGAGGAATATACTACGGATCCTA  320
BB2     TACATACATGCRAATGGAGCCTCAATATTCTTCATCTGCCTGTACATGCATGTAGGACGAGGAATATACTACGGATCCTA  320
IMP     TATATACATGCAAACGGAGCATCAATGTTTTTCATCTGTCTATTCATGCATGTAGGACGAGGCTTATACTACGGATCATA  320
COW     TACATACACGCAAACGGAGCTTCAATGTTTTTTATCTGCTTATATATGCACGTAGGACGAGGCTTATATTATGGGTCTTA  320
BUF     TACATGCACGCAAACGGAGCTTCAATATTCTTCATCTGCTTATATATACACGTGGGACGAGGCCTATACTATGGATCCTA  320
NIL     TACATACACGCAAATGGAGCTTCAATATTCTTTATTTGCCTGTACATACATGTAGGACGAGGACTATATTATGGATCATA  320

                330       340       350       360       370       380       390       400
                .         .         .         .         .         .         .         .
GE1     CACCTTCCTAGAAACATGAAACATCGGAGTAATTCTCCTATTCACGGTAATAGCCACAGCATTTATGGGATATGTACTAC  400
CE1     TACTTTTCTAGAAACATGAAACATTGGAGTAATTCTCTTATTCACAGTAATAGCCACAGCATTCATGGGCTATGTACTAC  400
CE2     TACTTTTCTAGAAACATGAAACATTGGAGTAATTCTCTTATTCACAGTAATAGCCGCAGCATTCATGGGCTATGTACTAC  400
GK1     CACTTTTCTAGAGACATGAAACATCGGAGTAATTCTCCTATTCACAGTTATAGCCACAGCATTCATAGGCTACGTACTAC  400
GK2     CACTTTTCTAGAGACATGAAACATCGGAGTAATTCTCCTATTCACAGTTATAGCCACAGCATTCATAGGCTACGTGCTAC  400
LK1     CACTTTCCTAGAAACATGAAATATCGGAGTAATTCTTCTATTCACAGTAATAGCCACAGCGTTCATAGGATACGTACTAC  400
CN1     CACTTTCCTAGAGACATGAAACGTCGGGGTAATTCTTCTATTTATAGTAATAGCCACAGCATTCATAGGATACGTGCTGC  400
CN2     CACTTTCCTAGAGACATGAAACGTCGGGGTAATTCTTCTATTTATAGTAATAGCCACAGCATTCATAGGATACGTGCTGC  400
MN1     CACCTTCCTAGAGACATGAAATATCGGAGTAATCCTCCTATTTACAGTAATAGCCACAGCATTCATGGGTTACGTACTGC  400
MN2     CACCTTCCTAGAGACATGAAATATCGGAGTAATCCTCCTATTTACAGTAATAGCCACAGCATTCATGGGTTACGTACTGC  400
B1      CACCTTCCTAGAAACATGAAACATCGGAGTAATTCTCCTATTTACAGTAATAGCCACAGCATTCATGGGCTACGTACTAC  400
B2      CACCTTCCTAGAAACATGAAACATTGGAGTAATTCTCCTATTTACAGTAATAGCCACAGCATTCACGGGCTACGTACTAC  400
SIT1    CACCTTCCTAGAAACATGAAACATCGGAGTAATTCTCCTATTTACAGTAATAGCCACAGCATTCATGGGCTACGTACTAC  400
BB1     CACCTTCCTAGAAACATGAAACATCGGAGTAATTCTCCTATTCACAGTAATAGCCACAGCATTCATGGGCTACGTACTAC  400
BB2     CACCTTCCTAGAAACATGAAACATCGGAGTAATTCTCCTATTCACAGTAATAGCCACAGCATTCATGGGCTACGTACTAC  400
IMP     CTTATTTCTAGAAACATGAAACATTGGAATTATTCTTCTATTCGCAACAATAGCCACAGCATTCATAGGCTATGTCCTGC  400
COW     CACTTTTCTAGAAACATGAAATATCGGAGTAATCCTTCTGCTCACAGTAATAGCCACAGCATTCATAGGATACGTCCTAC  400
BUF     CACTTTTTTAGAAACATGAAACATCGGAGTAATCCTCCTATTCACAGTAATAGCTACGGCATTCATAGGATATGTACTGC  400
NIL     CACATTCCTAGAAACATGAAACATCGGAGTAATTCTCCTATTCACAGTAATAGCCACAGCATTCATAGGCTACGTACTAC  400

                410       420       430       440       450       460       470       480
                .         .         .         .         .         .         .         .
GE1     CATGAGGACAAATATCATTCTGAGGAGCAACAGTCATCACAAACCTCCTGTCAGCAATCCCTTATATTGGCACCAGCTTA  480
CE1     CATGAGGACAAATATCATTCTGAGGAGCAACAGTTATCACAAACCTTCTATCAGCAATCCCTTATATCGGCACCAGCCTA  480
CE2     CATGAGGGCAAATATCRTTCTGAGGAGCAACAGTTATCACAAACCTCCTATCAGCAATCCCTTATATCGGCACCAGCCTA  480
GK1     CATGAGGACAAATATCATTCTGAGGAGCAACAGTAATTACAAACCTTTTATCAGCAATCCCTTATATTGGCACCAACCTA  480
GK2     CATGAGGACAAATATCATTCTGAGGAGCAACAGTAATTACAAACCTTTTATCAGCAATCCCTTATATTGGCACCAACCTA  480
LK1     CCTGAGGGCAAATATCATTCTGAGGAGCAACAGTTATCACGAACCTCCTATCAGCAATCCCATACATTGGCACCAACCTA  480
CN1     CATGAGGACAAATATCATTCTGAGGAGCAACGGTCATCACAAACCTCCTATCAGCAATCCCATATATTGGCACCAACCTA  480
CN2     CATGAGGACAAATATCATTCTGAGGAGCAACGGTCATCACAAACCTCCTATCAGCAATCCCATATATTGGCACCAACCTA  480
MN1     CATGAGGACAAATATCATTCTGAGGAGCAACAGTCATCACAAACCTTTTATCAGCAATTCCTTAGATCGGCACCAACCTA  480
MN2     CATGAGGACAAATATCATTCTGAGGAGCAACAGTCATCACAAACCTTTTATCAGCAATTCCTTAGATCGGCACCAACCTA  480
B1      CATGAGGACAAATATCATTTTGAGGAGCAACAGTCATCACCAACCTTCTATCAGCAATCCCTTATATTGGCACCAGCCTA  480
B2      CATGAGGACAAATATCATTTTGAGGAGCAACAGTCATCACAAACCTTCTATCAGCAATCCCTTATATTGGCACCAGCCTA  480
SIT1    CATGAGGACAAATATCATTTTGAGGAGCAACAGTCATCACAAACCTTCTATCAGCTATCCCTTATATCGGCACCAGCCTA  480
BB1     CATGAGGACAAATATCATTTTGAGGAGCAACGGTAATCACAAACCTTCTATCGGCAATCCCTTATATTGGCACCAGCCTA  480
BB2     CATGAGGACAAATATCATTCTGAGGAGCAACGGTAATCACAAACCTCCTATCGGCAATCCCTTATATTGGCACCAGCCTA  480
IMP     CATGAGGACAAATATCATTCTGAGGAGCAACAGTCATTACAAATCTCCTCTCAGCAATCCCATACATTGGTACAAACCTA  480
COW     CATGAGGACAAATATCATTCTGAGGAGCAACAGTCATCACCAACCTCTTATCAGCAATCCCATACATCGGCACAAATTTA  480
BUF     CATGAGGACAAATATCATTCTGAGGGGCAACAGTAATCACTAATCTCCTCTCAGCAATCCCATATATCGGCACAAACCTA  480
NIL     CATGAGGACAAATATCATTTTGAGGAGCAACAGTTATTACCAATCTATTATCAGCAATCCCATACATCGGCACAAACCTA  480
```

```
                490       500       510       520       530       540       550       560
                .         .         .         .         .         .         .         .
GE1     GTTGAATGGATCTGAGGAGGCTTTTCAGTAGATAAAGCAACCCTAACCCGATTCTTCGCCTTCCACTTTATCCTCCCATT    560
CE1     GTTGAGTGAATCTGAGGAGGCTTTTCAGTAGACAAAGCAACCTTGACCCGATTCTTCGCCTTCCACTTCATCCTCCCATT    560
CE2     GTTGAGTGAATCTGAGGAGGCTTTTCAGTAGACAAAGCAACCTTAACCCGATTCTTCGCCTTCCACTTCATCCTCCCATT    560
GK1     GTTGAATGAATCTGAGGAGGCTTTTCAGTGGACAAAGCAACCTTAACCCGATTCTTCGCTTTCCACTTCATCCTCCCGTT    560
GK2     GTTGAATGAATCTGAGGAGGCTTTTCAGTGGACAAAGCAACCTTAACCCGATTCTTCGCTTTCCACTTCATCCTCCCGTT    560
LK1     GTCGAATGAATCTGAGGAGGGTTTTCAGTGGACAAAGCAACCCTAACCCGATTTTTCGCCTTCCACTTCATCCTTCCATT    560
CN1     GTTGAATGAATCTGAGGAGGCTTCTCGGTAGACAAGGCAACCCTAACCCGATTTTTCGCCTTCCACTTCATCCTCCCGTT    560
CN2     GTTGAATGAATCTGAGGAGGCTTCTCGGTAGACAAGGCAACCCTAACCCGATTTTTCGCCTTCCACTTCATCCTCCCGTT    560
MN1     GTCGAATGAATCTGAGGAGGCTTTTCAGTAGACAAAGCAACCTTGACCCGATTCTTCGCCTTCCACTTTATTCTTCCATT    560
MN2     GTCGAATGAATCTGAGGAGGCTTTTCAGTAGACAAAGCAACCTTGACCCGATTCTTCGCCTTCCACTTTATTCTTCCATT    560
B1      GTCGAATGAATCTGAGGGGGCTTTTCAGTAGACAAAGCAACCTTAACCCGATTCTTCGCCTTCCACTTTATCCTTCCATT    560
B2      GTCGAATGAATCTGAGGGGGCTTTTCAGTAGACAAAGCAACCTTAACCCGATTCTTCGCCTTCCACTTTATCCTTCCATT    560
SIT1    GTTGAGTGGATCTGAGGAGGCTTTTCAGTAGACAAAGCAACCCTGACCCGATTCTTCGCTTTCCACTTTATTTTCCCATT    560
BB1     GTCGAATGAATCTGAGGAGGCTTTTCAGTAGACAAGGCAACCTTAACCCGATTCTTCGCCTTCCACTTTATCCTCCCATT    560
BB2     GTCGAATGAATCTGAGGAGGCTTTTCAGTAGACAAGGCAACCTTAACCCGATTCTTCGCCTTCCACTTTATCCTCCCATT    560
IMP     GTAGAATGAATCTGAGGAGGGTTCTCAGTAGACAAAGCAACCCTTACCCGATTTTTCGCCTTCCACTTCATCTTTCCATT    560
COW     GTCGAATGAATCTGAGGCGGATTTTCAGTAGACAAAGCAACCCTTACCCGATTTTTCGCTTTCCACTTTATCCTTCCATT    560
BUF     GTCGAATGAATCTGAGGGGGATTCTCAGTAGATAAAGCAACCCTCACCCGATTCTTCGCTTTCCACTTCATCCTCCCATT    560
NIL     GTTGAATGAATCTGAGGCGGGTTCTCAGTAGACAAAGCAACCCTAACCCGATTCTTCGCTTTCCACTTTATCCTCCCATT    560

                570       580       590       600       610       620       630       640
                .         .         .         .         .         .         .         .
GE1     TATTATTGCAGCACTAGCTATTGTCCACCTACTATTCCTCCACGAAACAGGATCCAACAACCCAACAGGAATCTCATCAG    640
CE1     TATTATTGCAGCACTAGCCATGGTCCACCTACTATTCCTCCACGAAACAGGATCCAACAACCCAACAGGAATCTCATCAG    640
CE2     TATTATTGCAGCACTAGCCATGGTCCACCTACTATTCCTCCACGAAACAGGATCCAACAACCCAACAGGAATCTCATCAG    640
GK1     CATTATTGCAGCACTAGCCATAGTCCACCTACTATTTCTCCACGAAACAGGATCTAACAACCCAACAGGAATCTCATCAG    640
GK2     CATTATTGCAGCACTAGCCATAGTCCACCTACTATTTCTCCACGAAACAGGATCTAACAACCCAACAGGAATCTCATCAG    640
LK1     TATTATCGCAGCACTAGCCCTTGTCCACCTTCTATTCCTCCACGAAACAGGATCCAACAACCCAACAGGAATCTCGTCAG    640
CN1     TATTATTACAGCGCTGGTTATGGTCCACCTATTATTCCTCCATGAAACAGGATCCAACAACCCAACAGGAATCTCATCAG    640
CN2     TATTATTACAGCGCTGGTTATGGTCCACCTATTATTCCTCCATGAAACAGGATCCAACAACCCAACAGGAATCTCATCAG    640
MN1     CATTATTGCAGCACTAGCCCTGGTACACCTACTATTTCTCCACGAAACAGGATCCAACAACCCAACAGGAATCTCATCAG    640
MN2     CATTATTGCAGCACTAGCCCTGGTACACCTACTATTTCTCCACGAAACAGGATCCAACAACCCAACAGGAATCTCATCAG    640
B1      TATTATTACAGCACTAGCCATGGTACACCTACTATTCCTCCACGAAACAGGATCCAACAACCCAACAGGAATCTCATCAA    640
B2      TATTATTACAGCACTAGCCATGGTACACCTACTATTCCTCCACGAAACAGGATCCAACAACCCAACAGGRATCTCATCRA    640
SIT1    TATTATTGCAGCACTAGCCATGGTACACCTACTATTTCTCCACGAGACAGGATCCAACAACCCAACAGGAATCTCATCAG    640
BB1     CATTATCGCAGCACTAGCCATGGTACACCTACTATTTCTCCACGAAACAGGATCCAACAACCCAACAGGAATCCCATCAG    640
BB2     CATTATCGCAGCACTAGCCATAGTACACCTACTATTTCTCCACGAAACAGGATCCAACAACCCAACAGGAATCCCATCAG    640
IMP     CATCATTGCAGCACTCGCCATAGTTCACCTACTTTTTCTCCACGAAACAGGATCTAACAACCCCACAGGAATCTCATCAG    640
COW     TATCATCATAGCAATTGCCATAGTCCACCTACTATTCCTCCACGAAACAGGCTCCAACAATCCAACAGGAATCTCCTCAG    640
BUF     CATTATTGCAGCACTTGCAATAATCCACCTATTATTTCTCCACGAAACAGGATCCAACAACCCAACAGGAATCTCATCAG    640
NIL     CATCATTGCAGCCCTCGCAATAATCCATCTACTCTTCCTCCATGAAACAGGGTCTAACAATCCAACAGGAATTTCATCAG    640

                650       660       670       680       690       700       710       720
                .         .         .         .         .         .         .         .
GE1     ACATAGACAAAATCCCATTCCACCCTTACTACACTATCAAGGACATCCTAGGCGCCCTACTACTAATCCTAGCCCTAATA    720
CE1     ACACAGACAAAATCCCATTCCACCCTTACCACACTATCAAGGACATTCTAGGCGCCCTACTACTAATCCTAACTCTAATA    720
CE2     ACACAGACAAAATCCCATTCCACCCTTACCACACTATCAAGGACATTCTAGGCGCCCTACTACTAATCCTAACTCTAATA    720
GK1     ACATAGACAAAATCCCATTCCACCCTTACTATACTATCAAAGATATCTTAGGCGCCCTACTGTTAGTCCTAGCTCTAATA    720
GK2     ACATAGACAAAATCCCATTCCACCCTTACTATACTATCAAAGATATCTTAGGCGCCCTACTGTTAGTCCTAGCTCTAATA    720
LK1     ACACAGACAAAATCCCATTCCACCCCTACTATACAATTAAAGACATTCTAGGTGCCCTATTACTAATTCTAGCTCTAATA    720
CN1     ACATAGACAAAATTCCATTCCACCCTTATTACACTATCAAGGACATCCTAGGCGCCCTACTATTAATCCTAGCCCTAATA    720
CN2     ACATAGACAAAATTCCATTCCACCCTTATTACACTATCAAGGACATCCTAGGCGCCCTACTATTAATCCTAGCCCTAATA    720
MN1     AMATAGACAAAATTCCATTCCACCCCTACTACACTATTAAGGATATCCTAGGCGCCCTTCTACTAATCCTAACTCTAATA    720
MN2     AMATAGACAAAATTCCATTCCACCCCTACTACACTATTAAGGATATCCTAGGCGCCCTTCTACTAATCCTAACTCTAATA    720
B1      ACATAGACAAAATTCCATTTCACCCTTACTACACTATTAAGGACATCCTAGGTGCCCTACTGCTAATCCTAACTCTAATA    720
B2      ACATAGACAAAATTCCATTTCACCCTTACTACACTATTAAGGACATCCTAGGTGCCCTACTGCTAATCCTAACTCTAATA    720
SIT1    ACATAGACAAAATCCCATTTCACCCTTACTACACTATCAAGGACATCTTAGGTGTCCTACTGTTAATCCTAACTCTAATA    720
BB1     ACATAGATAAAATTCCATTCCACCCTTACTACACCATTAAAGACATCCTAGGTGCCCTACTACTAATCCTAATTCTAATA    720
BB2     ACATAGATAAAATTCCATTCCACCCTTACTACACCATTAAAGACATCCTAGGTGCCCTACTACTAATCCTAATTCTAATA    720
IMP     ATTCAGATAAAATTCCATTCCACCCTTACTATACTATTAAAGACATCCTAGGAGTCCTACTAATAATTTTAGTCCTAATA    720
COW     ACGTAGACAAAATCCCATTCCACCCCTACTATACCATTAAGGACATCTTAGGGGCCCTCTTACTAATTCTAGCTCTAATA    720
BUF     ACACAGACAAAATCCCATTCCACCCCTACTACACCATCAAAGACATCCTAGGTGCTCTACTACTAATCCTAGCCCTAATA    720
NIL     ACGCAGATAAAATCCCATTTCACCCCTACTACACTATTAAAGACATTCTAGGAGCCCTACTACTTATTCTAGCCCTAATA    720
```

```
                730       740       750       760       770       780       790       800
                .         .         .         .         .         .         .         .
GE1     TTACTAGTACTATTCGCACCCGACCTCCTCGGAGACCCTGATAACTACACTCCAGCAAACCCACTCAGCACACCTCCCCA  800
CE1     CTTCTAGTACTATTCGCACCCGACCTCCTCGGAGACCCTGACAACTACACCCCAGCAAACCCACTCAACACACCTCCCCA  800
CE2     CTTCTAGTACTATTCGCACCCGACCTCCTCGGAGACCCTGACAACTACACCCCAGCAAACCCACTCAACACACCTCCCCA  800
GK1     CTACTAGTACTATTTACACCCGACCTCCTCGGAGACCCTGACAACTACACCCCAGCAAACCCACTCAACACACCTCCCCA  800
GK2     CTACTAGTACTATTTACACCCGACCTCCTCGGAGACCCTGACAACTACACCCCAGCAAACCCACTCAACACACCTCCCCA  800
LK1     CTTCTAGTGCTGTTTGCACCTGATCTCCTTGGAGACCCCGACAACTATACCCCAGCAAACCCACTCAACACACCCCCTCA  800
CN1     GTACTAGTACTATTCACACCTGACCTCCTCGGAGACCCCGACAACTACACCCCAGCGAACCCCCTCAATACACCTCCCCA  800
CN2     GTACTAGTACTATTCACACCTGACCTCCTCGGAGACCCCGACAACTACACCCCAGCGAACCCCCTCAATACACCTCCCCA  800
MN1     CTACTAGTACTATTCGCACCCGACCTCCTCGGAGACCCTGACAACTACACTCCAGCAAACCCACTCAACACACCTCCCCA  800
MN2     CTACTAGTACTATTCGCACCCGACCTCCTCGGAGACCCTGACAACTACACTCCAGCAAACCCACTCAACACACCTCCCCA  800
B1      CTCCTAGTACTATTCGCACCCGACCTTCTCGGAGACCCCGACAACTACACCCCAGCAAACCCACTCAACACACCACCTCA  800
B2      CTCCTAGTACTATTCGCACCCGACCTTCTCGGAGACCCCGACAACTACACCCCAGCAAACCCACTCAACACACCACCTCA  800
SIT1    CTCCTAGTACTATTCGCACCCGACCTCCTCGGAGACCCCGACAACTATACCCCAGCAAACCCACTCATCACACCACCTCA  800
BB1     CTCCTAGTATTATTCGCACCCGACCTCCTCGGAGACCCCGACAACTATACCCCAGCAAATCCACTCAACACACCACCCCA  800
BB2     CTCCTAGTATTATTCGCACCCGACCTCCTCGGAGACCCCGACAACTATACCCCAGCAAATCCACTCAACACACCACCCCA  800
IMP     CTCCTAGTACTATTCACACCCGACCTACTAGGAGACCCAGACAACTACACCCCCGCAAACCCACTCAACACACCTCCCCA  800
COW     CTACTAGTACTATTTGCGCCCAACCTCCTCGGAGACCCAGATAACTACACCCCAGCCAATCCACTCAACACACCTCCTCA  800
BUF     CTATTAGTACTATTCTCACCCGATCTCCTCGGAGACCCAGACAACTACACCCCAGCAAATCCACTCAACACACCCCCTCA  800
NIL     ATACTAGTACTATTCGCACCCGACCTCCTCGGAGACCCAGACAACTACACCCCAGCAAACCCACTTAGCACACCTCCCCA  800

                810       820       830       840       850       860       870       880
                .         .         .         .         .         .         .         .
GE1     TATCAAACCCGAATGATACTTTCTATTCGCATATGCAATCTTACGATTAATTCCCAACAAACTAGGGGGAGTTCTAGCTC  880
CE1     TATTAAACCCGAATGGTACTTCCTATTCGCATATGCAATCCTACGATCAATCCCCAACAAACTAGGAGGGGTTCTAGCTC  880
CE2     TATTAAACCCGAATGGTACTTCCTATTCGCATATGCAATCCTACGATCAATCCCCAACAAACTAGGAGGGGTTCTAGCTC  880
GK1     TATTAAACCCGAATGATATTTCCTATTCGCATACGCAATCCTACGATCGATCCCCAACAAGCTCGGAGGAGTTCTAGCTC  880
GK2     TATTAAACCCGAATGATATTTCCTATTCGCATACGCAATCCTACGATCGATCCCCAACAAGCTCGGAGGAGTTCTAGCTC  880
LK1     CATTAAACCCGAATGGTACTTCCTATTCGCATACGCAATTCTGCGATCAATCCCAAATAAACTAGGAGGGGTCCTAGCTC  880
CN1     TATCAAACCTGAATGATATTTCCTGTTCGCATATGCAATCCTACGATCTATCCCCAACAAGCTAGGAGGTGTCCTAGCCC  880
CN2     TATCAAACCTGAATGATATTTCCTGTTCGCATATGCAATCCTACGATCTATCCCCAACAAGCTAGGAGGTGTCCTAGCCC  880
MN1     TATCAAACCCGAATGATACTTCCTATTCGCATATGCAATCCTACGATCAATCCCCAACAAACTAGGAGGAGTTCTAGCTC  880
MN2     TATCAAACCCGAATGATACTTCCTATTCGCATATGCAATCCTACGATCAATCCCCAACAAACTAGGAGGAGTTCTAGCTC  880
B1      CATCAAACCTGAATGATACTTCCTATTCGCATATGCAATCCTACGATCAATCCCTAATAAACTAGGAGGGGTTCTAGCCC  880
B2      CATCAAACCTGAATGATACTTCCTATTCGCATATGCAATCCTACGATCAATCCCTAATAAACTAGGAGGGGTTCTAGCCC  880
SIT1    CATCAAGCCCGAATGGTATTTCCTATTTGCATATGCAATCCTACGATCAATCCCTAATAAACTAGGAGGAGTTCTAGCCC  880
BB1     CATCAAACCCGAATGGTATTTCCTATTCGCGTATGCAATCCTACGATCAATTCCCAACAAACTAGGAGGGGTCCTAGCCC  880
BB2     CATCAAACCCGAATGGTATTTCCTATTCGCATATGCAATCCTACGATCAATCCCCAACAAACTAGGAGGAGTCCTAGCCC  880
IMP     CATCAAGCCCGAATGGTACTTCCTATTCGCATACGCAGTCCTACGATCAATCCCCAACAAGCTAGGAGGAGTCCTAGCCT  880
COW     CATCAAACCCGAGTGGTACTTCCTATTTGCATATGCAATTCTACGATCAATCCCCAATAAACTAGGAGGAGTCCTAGCCC  880
BUF     TATCAAACCCGAATGATATTTCCTATTCGCATACGCGATCTTACGATCGATTCCCAACAAACTAGGGGGAGTTTTAGCCC  880
NIL     TATTAAGCCCGAATGGTATTTCCTGTTCGCATACGCAATTCTACGATCAATCCCCAACAAACTAGGAGGAGTTATAGCCC  880

                890       900       910       920       930       940       950       960
                .         .         .         .         .         .         .         .
GE1     TAGTACTCTCAATCCTAGTTCTCATACTAATACCCCTACTACACACATCCAAACAACGAAGCATAATATTCCGACCATTA  960
CE1     TAGTACTCTCAATTCTAATCCTTATCTTAATACCCCTACTACACACATCCAAACAACGAAGCATAATGTTCCGACCATTA  960
CE2     TAGTACTCTCAATTCTAATCCTTATCTTAATACCCCTACTACACACATCCAAACAACGAAGCATAATGTTCCGACCATTA  960
GK1     TAGTACTATCAATTCTTATTCTTATCTTTTTACCTTTACTACACACATCCAAACAACGAAGCATAATATTCCGACCATTA  960
GK2     TAGTACTATCAATTCTTATTCTTATCTTTTTACCTTTACTACACACATCCAAACAACGAAGCATAATATTCCGACCATTA  960
LK1     TAATACTCTCAATCCTAATCCTTATCTTAATACCTCTACTGCATGCATCCAAACAACGAAGCATAATATTCCGACCACTC  960
CN1     TAGTACTCTCAATCCTAATTCTAATCCTAATACCCCTACTACACATATCAAAACAACGAAGTATAATATTCCGACCATTA  960
CN2     TAGTACTCTCAATCCTAATTCTAATCCTAATACCCCTACTACACATATCAAAACAACGAAGTATAATATTCCGACCATTA  960
MN1     TAGTACTCTCAATCCTAATCCTCATCCTAATACCCCTGTTACATACATCCAAACAACGAAGCATGATATTCCGACCACTA  960
MN2     TAGTACTCTCAATCCTAATCCTCATCCTAATACCCCTGTTACATACATCCAAACAACGAAGCATGATATTCCGACCACTA  960
B1      TAGTACTCTCAATCCTAATCCTCATTCTAATACCCCTACTACATGTATCCAAACAACGAAGCATAATATTCCGACCATTA  960
B2      TAGTACTCTCAATCCTAATCCTCATTCTAATACCCCTACTACATGTATCCAAACAACGAAGCATAATATTCCGACCATTA  960
SIT1    TAGTACTCTCAATCCTTATCCTCATCTTAATGCCCCTACTACATGTATCCAAACAACGAAGCATAATATTTCGACCATTA  960
BB1     TAGTACTCTCAATCCTAATTCTCATCCTAATACCCCTACTACATACATCCAAACAACGAAGCATAATATTTCGACCATTA  960
BB2     TAGTACTCTCAATCCTAATTCTCATCCTAATACCCCTACTACATACATCCAAACAACGAAGCATAATATTTCGACCATTA  960
IMP     TAGTTCTTTCAATCCTAATCCTAGCCCTCATACCAATACTCCACATATCTAAACAACGAAGCATAATATTCCGGCCAATT  960
COW     TAGTCTTCTCCATCCTAATCCTTATTCTCATTCCCTTACTACACACATCCAAACAGCGAAGCATAATATTCCGACTGCTC  960
BUF     TAGTCCTCTCCATCCTAATCCTCATTATTATACCCCTGCTACACACATCCAAACAACGAAGCATGATATTCCGACCACTC  960
NIL     TAGTCCTCTCCATCCTAATCCTCATTCTCATACCCCTTCTACACACATCAAAGCAACGAAGCATAATATTCCGACCACTC  960
```

```
                970       980       990       1000      1010      1020      1030      1040
                .         .         .         .         .         .         .         .
GE1     AGCCAATGTTTCTTCTGAATCTTAGCAGCAGACCTACTAACACTTACATGAATCGGAGGACAACTAGTTGAACACCCATA    1040
CE1     AGTCAATGCCTCTTCTGAGTCCTAGCAGCAGACCTACTAACACTCACATGAATCGGAGGACAACCAGTTGAACACCCATA    1040
CE2     AGTCAATGCCTCTTCTGAGTCCTAGCAGCAGACCTACTAACACTCACATGAATCGGAGGACAACCAGTTGAACACCCATA    1040
GK1     AGCCAATGCCTCTTCTGAATCTTAGTAGCAGATCTACTAACACTAACATGAATCGGAGGACAACCAGTTGAACACCCATA    1040
GK2     AGCCAATGCCTCTTCTGAATCTTAGTAGCAGATCTACTAACACTAACATGAATCGGAGGACAACCAGTTGAACACCCATA    1040
LK1     AGCCAATGCCTATTCTGAATCTTAGTAGCAGACCTACTAACACTCACATGAATTGGAGGGCAACCAGTCGAACACCCATA    1040
CN1     AGCCAATGCTTATTCTGACTCTTAGTGGCAGACCTACTAACCCTTACATGAATTGGAGGACAGCCAGTTGAACACCCATA    1040
CN2     AGCCAATGCTTATTCTGACTCTTAGTGGCAGACCTACTAACCCTTACATGAATTGGAGGACAGCCAGTTGAACACCCATA    1040
MN1     AGCCAGTGCCTCTTCTGAATCTTAGCGGCAGACCTACTAACACTCACATGAATCGGAGGACAACCAGTCGAACATCCATA    1040
MN2     AGCCAGTGCCTCTTCTGAATCTTAGCGGCAGACCTACTAACACTCACATGAATCGGAGGACAACCAGTCGAACATCCATA    1040
B1      AGCCAATGCCTCTTCTGAATCTTAGCAGCAGACCTACTAACACTCACATGAATCGGAGGACAACCAGTCGAACACCCGTA    1040
B2      AGCCAATGCCTCTTCTGAATCTTAGCAGCAGACCTACTAACACTCACATGAATCGGAGGACAACCAGTCGAACACCCATA    1040
SIT1    AGCCAATGCCTCTTCTGAATCTTAGCAGCAGACCTACTAACACTCACATGAATTGGAGGACAACCGGTTGAACACCCATA    1040
BB1     AGCCAATGCCTCTTCTGAATCTTAGCAGCAGACCTACTAACACTCACATGAATCGGAGGACAACCAGTTGAGCACCCATA    1040
BB2     AGCCAATGCCTCTTCTGAATCTTAGCAGCAGACCTACTAACACTCACATGAATCGGAGGACAACCAGTTGAGCACCCATA    1040
IMP     AGCCAATGCTTATTCTGAATCCTAGTAGCTGACCTGCTAACATTCACATGAATTGGAGGACAACCAGTTGAACACCCATA    1040
COW     AGCCAATGCCTGTTCTGAATTTTAGTAGCAGACCTACTAACACTCACATGAATTGGAGGACAACCAGTTGAACATCCATA    1040
BUF     AGCCAATGCTTATTCTGAATCCTAGTAGCAGACCTGCTAACACTCACATGAATTGGAGGACAGCCGGTCGAACACCCATA    1040
NIL     AGCCAATGCATATTCTGAATTCTAGTAGCAAACCTACTAACGCTCACATGAATCGGAGGACAACCAGTCGAACACCCATA    1040

                1050      1060      1070      1080      1090      1100      1110      1120
                .         .         .         .         .         .         .         .
GE1     CATCATCATCGGACAACTAGCATCTATCATATATTTCCTCCTTATCCTAGTACTAATGCCAGTGGCTAGTATAATCGAAA    1120
CE1     CATCATCATCGGACAACTAGCATCCATCATATACTTCCTCCTCATCCTAGTGCTAATACCAGTAGCCAGCATAATCGAAA    1120
CE2     CATCATCATCGGACAACTAGCATCCATCATATACTTCCTCCTCATCCTAGTGCTAATACCAGTAGCCAGCATAATCGAAA    1120
GK1     TATAATCATCGGACAGCTAGCATCCATCATGTATTTTCTCCTTATCTTAGTACTAATACCAGTGACCAGCATAATCGAAA    1120
GK2     TATAATCATCGGACAGCTAGCATCCATCATGTATTTTCTCCTTATCTTAGTACTAATACCAGTGACCAGCATAATCGAAA    1120
LK1     TATTATTATCGGACAACTAGCATCTATCATATACTTCCTCCTTATCCTAGTGCTAATACCAGTGGCCAGTACAATCGAAA    1120
CN1     TATTATTATCGGACAACTAGCATCCATCATTTACTTCCTCCTTATTCTTGTACTAATACCAGTGATCAGTACAATCGAAA    1120
CN2     TATTATTATCGGACAACTAGCATCCATCATTTACTTCCTCCTTATTCTTGTACTAATACCAGTGATCAGTACAATCGAAA    1120
MN1     TATTATCATCGGACAGCTAGCATCCATCATATACTTCTTCCTTATTCTAGTATTAATACCAATAACCAGTATAATCGAAA    1120
MN2     TATTATCATCGGACAGCTAGCATCCATCATATACTTCTTCCTTATTCTAGTATTAATACCAATAACCAGTATAATCGAAA    1120
B1      TATTATCATTGGACAACTAGCATCCATCATATACTTCCTCATTATCCTAGTACTAATACCAGTAACCAGTATAATTGAAA    1120
B2      TATTATCATTGGACAACTAGCATCCATCATATACTTCCTCATTATCCTAGTACTAATACCAGTAACCAGTATAATTGAAA    1120
SIT1    TATTATTATTGGACAACTAGCATCCATCATATACTTCCTCATTATTCTAGTACTAATACCAGCAACCAGCATAATTGAAA    1120
BB1     TATCATCATTGGACAACTAGCATCAATCATATACTTCCTCATTATTCTAGTATTAATACCAGCAGCCAGTATAATTGAAA    1120
BB2     TATCATCATTGGACAACTAGCATCAATCATATACTTCCTCATTATTCTAGTATTAATACCAGCAGCCAGTATAATTGAAA    1120
IMP     CATTATTATTGGCCAACTAGCATCTATTATATATTTCACCCTTATCCTAGTGCTAATACCAGTAGCTAGCACCATCGAAA    1120
COW     CATCATCATTGGACAACTAGCATCTATCATATACTTTCTCCTCATCCTAGTGCTAATACCAACAGCCGGCACAGTTGAAA    1120
BUF     TATTATCATTGGACAACTAGCATCTATTATATATTTCCTCCTCATCCTAGTACTAATACCAACAGCTAGCACGATCGAGA    1120
NIL     CATCATTATTGGACAGCTAGCATCTATTATATACTTCCTCCTCATCCTGGTACTAATACCAACAGCCAGCATAATCGAAA    1120

                1130      1140      1150      1160      1170      1180      1190      1200
                .         .         .         .         .         .         .         .
GE1     ATAATCTCTTAAAATGAAGA??????????????????????????????????????????????????????????    1200
CE1     ACAACTTCCTAAAATGAAGAATGACACACCAAACACATGCTTACCACATAGTAAACCCAAGTCCTTGACCCCTTACGGGA    1200
CE2     ACAACTTCCTAAAATGAAGA??????????????????????????????????????????????????????????    1200
GK1     ACAACCTCCTAAAATGAAGAATGACACACCAAACACACGCTTATCACATAGTAAACCCAAGTCCTTGACCCCTTACGGGA    1200
GK2     ACAACTTCCTAAAATGAAGA??????????????????????????????????????????????????????????    1200
LK1     ATAACTTCCTAAAATGAAGAATGACACACCAAACACATGCTTACCACATAGTAAACCCAAGCCCTTGACCCCTTACAGGA    1200
CN1     ACAACCTCCTAAAATGAAGAATGACACACCAAACACATGCTTATCACATAGTAAACCCAAGCCCTTGACCCCTTACGGGA    1200
CN2     ACAACCTCCTAAAATGAAGA??????????????????????????????????????????????????????????    1200
MN1     ATAACTTTTTAAAATGAAGAATGACACACCAAACACATGCCTATCACATAGTAAACCCAAGTCCTTGACCCCTTACGGGA    1200
MN2     ATAACTTTTTAAAATGAAGA??????????????????????????????????????????????????????????    1200
B1      ATAACTTTCTAAAATGAAGAATGACACACCAAACACATGCTTATCACATGGTAAACCCAAGCCCTTGACCCCTTACAGGA    1200
B2      ATAACTTTCTAAAATGAAGA??????????????????????????????????????????????????????????    1200
SIT1    ATAACTTTTTAAAATGAAGAATGACACACCAAACACATGCTTATCACATGGTAAACCCAAGTCCTTGACCCCTTACAGGA    1200
BB1     ATAACTTTCTAAAATGAAGGATGACACACCAAACACATGCTTATCACATAGTAAACCCAAGTCCTTGACCCCTTACGGGA    1200
BB2     ATAGCTTTCTAAAATGAAGG??????????????????????????????????????????????????????????    1200
IMP     ACAACCTCCTAAAATGAAGAATGACACACCAAACCCATGCTTATCACATAGTAAACCCAAGCCCCTGACCCCTCACAGGA    1200
COW     ACAACCTACTAAAATGAAGAATGACACACCAAACTCATGCTTATCATATAGTAAACCCAAGCCCTTGACCTCTTACAGGA    1200
BUF     ACAATCTCTTAAAATGAAGAATGACACACCAAACCCATGCTTACCACATAGTAAACCCAAGCCCTTGACCTCTTACAGGA    1200
NIL     ATAATCTCTTAAAATGAAGA??????????????????????????????????????????????????????????    1200
```

```
                1210       1220       1230       1240       1250       1260       1270       1280
                   .          .          .          .          .          .          .          .
GE1     ????????????????????????????????????????????????????????????????????????????????????????????????    1280
CE1     GCCTTATCTGCTCTATTAATAACATTCGGCCTAATCATATGATTTCACTTCAACTCAACAGCCCTATTAATGCTTGGCTT                    1280
CE2     ????????????????????????????????????????????????????????????????????????????????????????????????    1280
GK1     GCCCTATCCGCTCTGCTAATAACATCTGGCCTAACCATGTGATTCCACTTCAACTCAACAATTCTACTAATACTTGGCCT                    1280
GK2     ????????????????????????????????????????????????????????????????????????????????????????????????    1280
LK1     GCCCTATCCGCCCTACTAATAACATCAGGCCTAACCATGTGATTCCACTACAACTCAACAATTCTCCTAATGCTCGGCCT                    1280
CN1     GCCCTATCCGCCCTACTAATAACATCTGGCCTAACCATATGATTCCACTTTAACTCAAAAATTCTATTAATGCTAGGCCT                    1280
CN2     ????????????????????????????????????????????????????????????????????????????????????????????????    1280
MN1     GCCCTATCCGCCCTATTAATAACATCTGGCCTAACCATATGATTCCACTTCAACTCAACAATTCTACTATTACTTGGCTT                    1280
MN2     ????????????????????????????????????????????????????????????????????????????????????????????????    1280
B1      GCCTTATCCGCTCTACTGATAACATCCGGTCTAACCATATGATTCCACTTCAACTCAATAATTCTATTAATACTTGGCTT                    1280
B2      ????????????????????????????????????????????????????????????????????????????????????????????????    1280
SIT1    GCCTTATCTGCCCTACTGATAACATCCGGCCTAACCATATGATTCCACTTCAACTCGATAATTCTATTAATACTTGGCTT                    1280
BB1     GCCCTATCTGCTCTACTAATAACATCCGGCCTAACCATGTGATTTCACTTCAACTCAATAATTCTATTGATACTTGGCTT                    1280
BB2     ????????????????????????????????????????????????????????????????????????????????????????????????    1280
IMP     GCATTATCCGCCCTCTTAATA-CATCCGGTTTAATTATATGATTTCACTTTAACTCAACAATCCTACTAATACTCGGCTT                    1279
COW     GCTTTGTCTGCCCTCTTAATAACATCCGGCCTAACCATGTGATTTCACTTTAACTCAATGACCCTGCTAATAATTGGCCT                    1280
BUF     GCCCTATCCGCCCTATTAATAACATCCGGTTTAATCATATGATTTCACTTTAACTCAAC-ACCCTGTTAATACTTGGCTT                    1279
NIL     ????????????????????????????????????????????????????????????????????????????????????????????????    1280

                1290       1300       1310       1320       1330       1340       1350       1360
                   .          .          .          .          .          .          .          .
GE1     ????????????????????????????????????????????????????????????????????????????????????????????????    1360
CE1     AACAACAAACATGCTTACAATATACCAATGGTGACGAGACATCATCCGAGAGAGCACCTTTCAAGGACACCACACCCCAG                    1360
CE2     ????????????????????????????????????????????????????????????????????????????????????????????????    1360
GK1     AACAACCAACATACTTACAATATACCAATGATGACGAGACATCATCCGAGAAAGCACCTTTCAAGGACACCACACCCCAG                    1360
GK2     ????????????????????????????????????????????????????????????????????????????????????????????????    1360
LK1     AACAACTAATATACTGACAATATATCAATGATGACGAGATATTATCCGAGAGAGCACCTTTCAAGGACATCACACTCCAA                    1360
CN1     AACAACTAATATGCTCACAATATACCAATGATGGCGAGACATCGTCCGAGAAAGTACCTTTCAAGGACACCACACCCCAA                    1360
CN2     ????????????????????????????????????????????????????????????????????????????????????????????????    1360
MN1     AACAACTAATATACTTACAATATACCAATGATGACGAGACATCATCCGAGAGAGCACCTTTCAAGGACACCACACCCCAG                    1360
MN2     ????????????????????????????????????????????????????????????????????????????????????????????????    1360
B1      AACGACTAACATACTTACAATATACCAATGATGACGAGACATCATTCGAGAGAGCACCTTTCAAGGACACCACACCCCAG                    1360
B2      ????????????????????????????????????????????????????????????????????????????????????????????????    1360
SIT1    AACAACTAACATACTTACAATATATCAATGATGACGAGACATCATTCGAGAAAGCACCTTTCAAGGACACCACACTCCAG                    1360
BB1     GACAACTAACATACTTACAATGTACCAATGATGACGAGACATCATCCGAGAGAGCACCTTTCAAGGACACCACACCCCAG                    1360
BB2     ????????????????????????????????????????????????????????????????????????????????????????????????    1360
IMP     AACAACAAATATACTCACAA-ATATCAATGATGACGAGATATTATCCGAGAAAGCACTTTCCAAGGCCATCACACTCCAA                    1358
COW     AACAACAAATATACTAACAATATACCAATGATGACGAGATGTTATCCGAGAAAGCACCTTCCAAGGGCACCATACCCCAG                    1360
BUF     AACAACAAATATACTTACAATATATCAATGGTGACGAGATATTATCCGAGAAAGCACTTTCCAAGGACACCACACCCCAG                    1359
NIL     ????????????????????????????????????????????????????????????????????????????????????????????????    1360

                1370       1380       1390       1400       1410       1420       1430       1440
                   .          .          .          .          .          .          .          .
GE1     ????????????????????????????????????????????????????????????????????????????????????????????????    1440
CE1     TCGTCCAAAAAGGTCTCCGCTACGGAATAATCCTTTTCATTATTTCTGAAGTCTTATTCTTCACCGGATTTTTCTGAGCA                    1440
CE2     ????????????????????????????????????????????????????????????????????????????????????????????????    1440
GK1     TCGTCCAAAAAGGTCTCCGCTATGGGATGATCCTTTTTATTATTTCTGAAGTCCTATTCTTCACCGGATTTTTCTGAGCG                    1440
GK2     ????????????????????????????????????????????????????????????????????????????????????????????????    1440
LK1     CTGTCCAAAAAGGCCTTCGCTACGGAATGATCCTTTTTATTATTTCTGAAGTCCTATTCTTCACCGGGTTCTTCTGAGCA                    1440
CN1     CTGTCCAAAAAGGCCTCCGCTATGGAATAATCCTATTTATTGTCTCTGAAGTCCTATTCTTCACCGGATTTTTCTGGGCA                    1440
CN2     ????????????????????????????????????????????????????????????????????????????????????????????????    1440
MN1     TCGTCCAAAAAGGCCTCCGCTACGGAATAATCCTTTTTATCATTTCTGAAGTCCTATTCTTCACCGGATTTTTCTGAGCA                    1440
MN2     ????????????????????????????????????????????????????????????????????????????????????????????????    1440
B1      TCGTTCAAAAGGGCCTCCGCTACGGAATAATCCTTTTTATTATCTCTGAAGTCCTATTCTTCACCGGATTCTTCTGAGCA                    1440
B2      ????????????????????????????????????????????????????????????????????????????????????????????????    1440
SIT1    TCGTCCAAAAAGGCCTCCGCTACGGAATAATCCTTTTTATTATTTCTGAAGTCCTATTCTTCACTGGATTCTTCTGAGCA                    1440
BB1     TCGTCCAAAAAGGTCTCCGCTACGGAATGATCCTTTTTATCATCTCTGAAGTCCTATTCTTCACTGGATTCTTCTGAGCA                    1440
BB2     ????????????????????????????????????????????????????????????????????????????????????????????????    1440
IMP     CCGTCCAAAA-GGTCTTCGCTATGGAATAATTCTCTTCATTATCTCTGAGGTCCTATTCTTCACTGGGTTTTTCTGAGCA                    1437
COW     CTGTCCAAAAAGGCCTCCGTTATGGAATAATTCTTTTTATTATCTCCGAAGTACTATTCTTTACCGGATTTTTCTGAGCT                    1440
BUF     TTGTTCAAAAAGGCCTTCGCTATGGAATAATCCTCTTCATTATTTCCGAAGTCCTATTTTTTACCGGATTTTTCTGAGCA                    1439
NIL     ????????????????????????????????????????????????????????????????????????????????????????????????    1440
```

```
                1450       1460       1470       1480       1490       1500       1510       1520
                   .          .          .          .          .          .          .          .
GE1     ??????????????????????????????????????????????????????????????????????????????????????????????    1520
CE1     TTTTACCACTCAAGCCTTGCTCCCACACCCGAATTAGGAGGCTGCTGACCACCAACAGGCATTCACCCACTCAACCCCCT                    1520
CE2     ??????????????????????????????????????????????????????????????????????????????????????????????    1520
GK1     TTTTACCACTCAAGCCTTGCCCCCACACCCGAATTAGGAGGTTGCTGACCACCAACAGGCATCCACCCACTCAACCCCCT                    1520
GK2     ??????????????????????????????????????????????????????????????????????????????????????????????    1520
LK1     TTCTACCACTCAAGCCTTGCCCCCACGCCTGAATTAGGAGGCTGCTGACCCCCAACAGGCATTCACCCACTCAACCCATT                    1520
CN1     TTTTACCACTCAAGCCTTGCTCCCACACCTGAATTAGGAGGCTGCTGACCACCAACAGGCATTCATCCACTCAATCCCCT                    1520
CN2     ??????????????????????????????????????????????????????????????????????????????????????????????    1520
MN1     TTTTACCACTCAAGCCTTGCTCCTACACCCGAACTAGGAGGTTGCTGACCACCAACAGGCATTCACCCACTCAACCCCCT                    1520
MN2     ??????????????????????????????????????????????????????????????????????????????????????????????    1520
B1      TTTTACCACTCAAGCCTTGCTCCCACACCCGAATTAGGAGGTTGCTGACCACCAACAGGTATTCACCCACTCAACCCCTT                    1520
B2      ??????????????????????????????????????????????????????????????????????????????????????????????    1520
SIT1    TTTTACCACTCAAGCCTTGCTCCTACACCCGAATTAGGAGGTTGCTGGCCACCAACAGGTATCCACCCACTCGACCCCCT                    1520
BB1     TTTTACCACTCAAGCCTTGCTCCTACACCCGAATTAGGAGGTTGCTGACCACCAACAGGTATTCACCCACTCAACCCCCT                    1520
BB2     ??????????????????????????????????????????????????????????????????????????????????????????????    1520
IMP     TTTTACCACT-AAGCCTTGCCCCTACGCCTGAACTAGGAGGTTGCTGACCCCCAACAGGAATTCATCCACTCAACCCCTT                    1516
COW     TTCTACCACTCAAGCCTCGCCCCCACCCCTGAACTAGGCGGCTGCTGACCCCCAACAGGCATTCACCCACTAAACCCCCT                    1520
BUF     TTCTATCACTCAAGCCTCGCCCCTACACCCGAACTAGGCGGCTGCTGACCTCCAACAGGCATCAATCCTCTCAACCCCCT                    1519
NIL     ??????????????????????????????????????????????????????????????????????????????????????????????    1520

                1530       1540       1550       1560       1570       1580       1590       1600
                   .          .          .          .          .          .          .          .
GE1     ??????????????????????????????????????????????????????????????????????????????????????????????    1600
CE1     AGAAGTCCCACTACTTAACACCTCCGTCTTATTAGCCTCAGGAGTCTCCATTACCTGAGCTCATCATAGCCTAATAGAAG                    1600
CE2     ??????????????????????????????????????????????????????????????????????????????????????????????    1600
GK1     GGAAGTCCCACTACTTAACACCTCCGTCTTATTAGCCTCAGGAGTCTCCATTACCTGAGCTCATCATAGCCTAATAGAAG                    1600
GK2     ??????????????????????????????????????????????????????????????????????????????????????????????    1600
LK1     AGAAGTCCCACTACTTAACACTTCCGTTTTACTGGCTTCAGGAGTCTCTATTACTTGAGCTCACCATAGCCTAATAGAAG                    1600
CN1     AGAAGTCCCACTACTTAACACCTCCGTCTTACTAGCTTCAGGAGTCTCTATTACCTGGGCCCATCATAGCCTAATAGAAG                    1600
CN2     ??????????????????????????????????????????????????????????????????????????????????????????????    1600
MN1     AGAAGTCCCACTACTTAATACCTCCGTCTTATTAGCCTCAGGAGTCTCTATTACCTGGGCCCACCATAGCCTAATAGAAG                    1600
MN2     ??????????????????????????????????????????????????????????????????????????????????????????????    1600
B1      AGAGGTCCCACTACTTAATACTTCCGTCCTATTAGCTTCAGGGGTTTCCATCACCTGAGCCCATCATAGCCTAATAGAAG                    1600
B2      ??????????????????????????????????????????????????????????????????????????????????????????????    1600
SIT1    AGAGGTCCCACTACTTAACACTTCCGTCCTACTAGCTTCAGGGGTCTCTATTACCTGAGCCCATCATAGCCTAATAGAAG                    1600
BB1     GGAAGTCCCACTACTCAACACCTCTGTCTTATTAGCTTCAGG--TCTCTATCACCTGAGCCCATCATAGCCTAATAGAAG                    1598
BB2     ??????????????????????????????????????????????????????????????????????????????????????????????    1600
IMP     AGAAGTCCCA-TACTCAATACCTCCGTCCTTCTGGCTTCAGGAGTTTCTATTACTTGAGCCCATCACAGCCTCATAGAAG                    1595
COW     AGAAGTCCCACTGCTCAACACCTCTGTCCTATTGGCTTCCGGAGTTTCTATTACCTGAGCCCATCATAGTTTAATAGAAG                    1600
BUF     AGAAGTCCCACTACTCAACACCTCCGTCCTGTTAGCCTCAGGAGTCTCTATTACCTGGGCCCACCATAGCCTGATGGAAG                    1599
NIL     ??????????????????????????????????????????????????????????????????????????????????????????????    1600

                1610       1620       1630       1640       1650       1660       1670       1680
                   .          .          .          .          .          .          .          .
GE1     ??????????????????????????????????????????????????????????????????????????????????????????????    1680
CE1     GACACCGCAACCACATACTACAAGCCCTATTTATTACTATTGCACTAGGAGTATATTTTACACTGTTACAAGCCTCAGAA                    1680
CE2     ??????????????????????????????????????????????????????????????????????????????????????????????    1680
GK1     GACACCGCAACCACATACTACAAGCCCTATTTATTACTATTGCACTAGGAGTATATTTTACACTGTTACAAGCCTCAGAA                    1680
GK2     ??????????????????????????????????????????????????????????????????????????????????????????????    1680
LK1     GAGACCGCAACCACATACTACAAGCCCTATTTATCACTATCGCACTAGGGATCTATTTTACACTACTACAGGCCTCAGAA                    1680
CN1     GGAACCGCAACCACATACTACAAGCCCTATTTATTACCATTGCACTAGGAGTGTACTTTACGCTACTGCAAGCCTCAGAG                    1680
CN2     ??????????????????????????????????????????????????????????????????????????????????????????????    1680
MN1     GACACCGCAACCACATACTACAAGCCCTATTTATTACTATCGCGCTAGGAGTATATTTCACACTATTACAAGCCTCAGAG                    1680
MN2     ??????????????????????????????????????????????????????????????????????????????????????????????    1680
B1      GACACCGCAACCACGTACTACAAGCCCTGTTCATTACTATCGCACTAGGAGTATATTTCACACTACTACAAGCCTCAGAG                    1680
B2      ??????????????????????????????????????????????????????????????????????????????????????????????    1680
SIT1    GACACCGCAACCACATGCTACAAGCCCTGTTTATTACTATCGCACTAGGAGTATATTTCACACTACTACAAGCCTCAGAG                    1680
BB1     GACACCGTAACCATATACTACAAGCCCTGTTTATTACTATCGCACTAGGGGTATATTTCACACTACTACAAGCCTCAGAA                    1678
BB2     ??????????????????????????????????????????????????????????????????????????????????????????????    1680
IMP     GAAATCGCAAC-ACATATTACAAGCCCTATTTATTACCATTGCACTAGGCGTATACTTCACACTGCTACAAGCCTCAGAA                    1674
COW     GGGACCGAAAGCATATATTACAAGCCCTATTTATCACCATCACATTAGGAGTCTACTTCACACTACTACAAGCCTCAGAA                    1680
BUF     GTAACCGAAGCCACATGTTACAAGCCCTGTTTATTACCATCACACTAGGGGTATATTTCACACTACTACAAGCCTCAGAG                    1679
NIL     ??????????????????????????????????????????????????????????????????????????????????????????????    1680
```

```
                1690       1700       1710       1720       1730       1740       1750       1760
                   .          .          .          .          .          .          .          .
GE1     ??????????????????????????????????????????????????????????????????????????????????????????????????    1760
CE1     TACTATGAAGCACCCTTTACTATCTCAGATGGTGTATATGGCTCAACTTTCTTTGTAGCCACAGGCTTCCACGGCCTCCA                        1760
CE2     ??????????????????????????????????????????????????????????????????????????????????????????????????    1760
GK1     TACTATGAAGCACCCTTTACTATCTCAGATGGTGTATATGGCTCAACTTTCTTTGTAGCCACAGGCTTCCACGGCCTCCA                        1760
GK2     ??????????????????????????????????????????????????????????????????????????????????????????????????    1760
LK1     TATTATGAAGCACCCTTTACCATCTCAGACGGGGTATATGGATCAACCTTTTTCGTAGCTACAGGCTTCCATGGCCTCCA                        1760
CN1     TACTATGAAGCGCCATTCACTATCTCAGATGGAGTGTATGGCTCAACTTTCTTCGTGGCCACTGGTTTCCATGGTCTCCA                        1760
CN2     ??????????????????????????????????????????????????????????????????????????????????????????????????    1760
MN1     TACTATGAAGCACCCTTTACTATCTCAGACGGGGTATACGGCTCAACCTTCTTTGTAGCCACAGGCTTCCACGGCCTCCA                        1760
MN2     ??????????????????????????????????????????????????????????????????????????????????????????????????    1760
B1      TACTACGAAGCACCCTTTACCATCTCAGATGGAGTATACGGCTCAACCTTCTTTGTAACCACAGGATTCCACGGTCTCCA                        1760
B2      ??????????????????????????????????????????????????????????????????????????????????????????????????    1760
SIT1    TACTATGAAGCACCCTTTACCATTTCAGATGGGGTGTATGGCTCAACCTTTTTTGTAGCCACAGGATTCCACGGCCTCCA                        1760
BB1     TACTATGAAGCGCCCTTTACCATCTCAGATGGAGTATATGGCTCAACTTTCTTTGTAGCCACAGGCTTCCACGGCCTCCA                        1758
BB2     ??????????????????????????????????????????????????????????????????????????????????????????????????    1760
IMP     TACTACGAAG-ACCCTTTACTATTTCAGACGGAGTTTACGGCTCAACCTTTTTCGTAGCCACAGGTTTTCACGGCCTACA                        1753
COW     TACTATGAAGCACCTTTTACTATCTCCGACGGAGTTTACGGCTCAACTTTTTTTGTAGCCACAGGCTTCCACGGCCTCCA                        1760
BUF     TACTATGAGGCACCCTTCACTATCTCAGACGGAGTGTACGGCTCAACTTTCTTCGTAGCCACAGGCTTCCATGGCCTCCA                        1759
NIL     ??????????????????????????????????????????????????????????????????????????????????????????????????    1760

                1770       1780       1790       1800       1810       1820       1830       1840
                   .          .          .          .          .          .          .          .
GE1     ??????????????????????????????????????????????????????????????????????????????????????????????????    1840
CE1     CGTCATTATTGGATCTACATTCTTAATTGTCTGTTTTTTCCGCCAACTAAAATTCCACTTCACCTCTAGCCATCATTTTG                        1840
CE2     ??????????????????????????????????????????????????????????????????????????????????????????????????    1840
GK1     CGTCATTATTGGATCTACATTCTTAATTGTCTGTTTTTTC-GCCAACTAAAATTCCACTTCACCTCTAGTCATCATTTTG                        1839
GK2     ??????????????????????????????????????????????????????????????????????????????????????????????????    1840
LK1     CGTCATTATTGGATCTACATTCGTATGCTTTTTCCGCCAACTAAAGTTTCATTTTACCTCTAGCCACCACTTTG                              1840
CN1     CGTCATTATTGGATCCACATTCTTAATTGTCTGCTTTTTTCGCCAACTAAAATTCCACTTTACCTCTAGTCACCACTTTG                        1840
CN2     ??????????????????????????????????????????????????????????????????????????????????????????????????    1840
MN1     CGTCATCATTGGATCCACATTCCTAATCGTTTGCTTTTTCCGCCAACTAAAATTCCACTTCACCTCTAGCCATCATTTTG                        1840
MN2     ??????????????????????????????????????????????????????????????????????????????????????????????????    1840
B1      CGTCATTATTGGATCCACGTTCTTAATTGTCTGCTTTTTCCGCCAACTAAAATTTCACTTTACCTCTAGTCATCACTTTG                        1840
B2      ??????????????????????????????????????????????????????????????????????????????????????????????????    1840
SIT1    CGTCATTATTGGAACCACATTTTTAATTGTTTGCTTTTTCCGCCAACTAAAATTTCACTTTACCTCTAGTCACCACTTTG                        1840
BB1     TGTCATTATTGGATCCACATTCTTAATTGTCTGCTTTTTCCGTCAACTAAAATTCCACTTTACCTCTAGTCATCACTTTG                        1838
BB2     ??????????????????????????????????????????????????????????????????????????????????????????????????    1840
IMP     CGTCATCATTG-ATCTACCTTCCTAATCGTTTGTTTTTTTCGTCAACTAAAATTCCACTTCACTTCTAGCCACCACTTTG                        1832
COW     CGTCATCATTGGGTCCACCTTCTTAATTGTCTGCTTCTTCCGCCAATTAAAATTTCATTTTACTTCTAACCACCACTTCG                        1840
BUF     CGTGATCATTGGCTCTACTTTCCTAATTGTCTGTTTCTTCCGTCAACTAAAATTCCATTTCACCTCTAACCACCACTTCG                        1839
NIL     ??????????????????????????????????????????????????????????????????????????????????????????????????    1840

                1850       1860       1870       1880       1890       1900       1910       1920
                   .          .          .          .          .          .          .          .
GE1     ??????????????????????????????????????????????????????????????????????????????????????????????????    1920
CE1     GCTTTGAAGCCGCTGCCTGATATTGACACTTTGTAGATGTAGTATGACTTTTCTTATATGTATCTATCTATTGATGAGGC                        1920
CE2     ??????????????????????????????????????????????????????????????????????????????????????????????????    1920
GK1     GCTTTGAAGCCGCTGCCTGATATTGACACTTTGTAGATGTAGTGTGACTTTTCTTATATGTATCTATCTATTGATGAGGC                        1919
GK2     ??????????????????????????????????????????????????????????????????????????????????????????????????    1920
LK1     GCTTTGAAGCCGCTGCCTGATACTGACACTTCGTAGACGTAGTGTGACTTTTCTTATACGTATCTATTTATTGATGAGGC                        1920
CN1     GCTTTGAAGCCGCTGCTTGATATTGACACTTTGTAGACGTAGTATGACTTTTCCTATACGTATCTATCTATTGATGAGGC                        1920
CN2     ??????????????????????????????????????????????????????????????????????????????????????????????????    1920
MN1     GTTTTGAAGCTGCTGCCTGATATTGACACTTCGTAGATGTAGTATGACTTTTCCTGTACGTGTCTATCTATTGATGAGGC                        1920
MN2     ??????????????????????????????????????????????????????????????????????????????????????????????????    1920
B1      GCTTTGAAGCTGCTGCCTGATATTGACACTTCGTAGATGTAGTATGACTTTTCCTGTATGTATCCATCTATTGATGAGGC                        1920
B2      ??????????????????????????????????????????????????????????????????????????????????????????????????    1920
SIT1    GTTTTGAAGCCGCTGCCTGATATTGACACTTCGTAGATGTAGTATGACTTTTCCTGTATGTATCCATCTATTGATGAGGC                        1920
BB1     GCTTTGAAGCTGCTGCCTGATATTGACACTTTGTAGATGTAGTATGACTTTTCCTATATGTATCCATCTACTGATGAGGC                        1918
BB2     ??????????????????????????????????????????????????????????????????????????????????????????????????    1920
IMP     GTTTCGAAGC-GCTGCCTGATATTGACACTTTGTAGACGTAGTATGACTTTTCCTTTACGTCTCTATCTATTGATGAGGC                        1911
COW     GCTTTGAAGCCGGTGCCTGATACTGACATTTCGTAGACGTAGTCTGACTTTTCCTCTATGTTTCTATCTATTGATGAGGC                        1920
BUF     GCTTTGAAGCCGCTGCCTGATACTGACACTTTGTAGACGTAGTCTGACTTTTCCTCTATGTATCTATCTATTGATGAGGT                        1919
NIL     ??????????????????????????????????????????????????????????????????????????????????????????????????    1920
```

```
                1930       1940       1950       1960       1970       1980       1990       2000
                   .          .          .          .          .          .          .          .
GE1     ??????????????????????????????????????????????????????????????????????????????????????????????????    2000
CE1     TCAGGAATTAGCGCCAAATCAAACAGATGAGTGAATTTTCCAAGTATTAATAGTACTTTCCTTGTGTATTATTGATACGG                        2000
CE2     ??????????????????????????????????????????????????????????????????????????????????????????????????    2000
GK1     TCAGGAATTAGCGCCAAATCAAACAGATGAGTGAATTTTCCAAGTATTAATAGTACTTTCCTTGTGTATTATTGATACGG                        1999
GK2     ??????????????????????????????????????????????????????????????????????????????????????????????????    2000
LK1     TCAGGAATTAGCGCCAAATCAAACAGATGAGTGAATTTTCCAAGTATTAATAGTACTTTCCTTGTGTATTATTGATATGG                        2000
CN1     TCAGGAATTAGTGCCAAATCAAACAGATGAGTGAATTTTCCAAGTATTAATAGTACTTTCCTTGTGTATTATTGATATGG                        2000
CN2     ??????????????????????????????????????????????????????????????????????????????????????????????????    2000
MN1     TCA??????????????????????????????????????????????????????????????????????????????????????????????    2000
MN2     ??????????????????????????????????????????????????????????????????????????????????????????????????    2000
B1      TCAGGAATTAGTGCCAAATCAAACAGATGAGTGAATTTTCCAAGTATTAATAGTACTTTCCTTGTGTATTATTGATACGG                        2000
B2      ??????????????????????????????????????????????????????????????????????????????????????????????????    2000
SIT1    TCA-----------------AAACAGATGAGTGAATTTTCCAAGTATTAATAGTACTTTCCTTGTGTATTATTGATACGG                        1983
BB1     TCAGGAATTAGCGCCAAATCAAACAGATGAGTGAATTTTCCAAGTATTAATAGTACTTTCCTTGTGTATTATTGATACGG                        1998
BB2     ??????????????????????????????????????????????????????????????????????????????????????????????????    2000
IMP     TCAGGAATTAGCACCAAATCAAACAGATGAGTGAATTTTCCAAGTATTAATAGTACTTTCCTTGTGTATTATTGATATGG                        1991
COW     TCCGGAATTAGCGCCAAATCAAACAGATGAGTGAATTTTCCAAGTATTAATAGTATTTTCCTTGTGTATTATTGATATGG                        2000
BUF     TCAGGAATTAGCGCCAAATCAAACAGATGAGTGAATTTTCCAAGTATTAATAGTATTTTCCTTGTGTATTATTGATATGG                        1999
NIL     ???GGAATTAGCACCAAATCAAACAGATGAGTGAATTTTCCCAGTATTAATAGTACTTTCCTTGTGTATTATTGATATGG                        2000

                2010       2020       2030       2040       2050       2060       2070       2080
                   .          .          .          .          .          .          .          .
GE1     ??????????????????????????????????????????????????????????????????????????????????????????????????    2080
CE1     CCTTACAAGTGTATCTGCTGAGTGCCACATAAAAACTAAAAAACTAAATTCATTTACAAACATGTAAATGTTTGCCATTG                        2080
CE2     ??????????????????????????????????????????????????????????????????????????????????????????????????    2080
GK1     CCTTACAAGTGTATCTGCTGAATGCCACATAAAAACTAAAAAACTAAATTCATTTACAAACATGTAAATGTTTGCCATTG                        2079
GK2     ??????????????????????????????????????????????????????????????????????????????????????????????????    2080
LK1     CCTTACAAGTGTATCTGCTGAATGCCACATAAAAACTAAAAAACTAAATTCATTTACAAACATGTAAATGTTTGCCATTG                        2080
CN1     CCTTACAAGTGTATCTGCTGAATGCCACATAAAAACTAAAAAACTAAATTCATTTACAAACATGTAAATGTTTGCCATTG                        2080
CN2     ??????????????????????????????????????????????????????????????????????????????????????????????????    2080
MN1     ??????????????????????????????????????????????????????????????????????????????????????????????????    2080
MN2     ??????????????????????????????????????????????????????????????????????????????????????????????????    2080
B1      CCTTACAAGTGTATCTGCTGAATGCCACATAAAAACTAAAAAACCTAAATTCATTTACAAACATGTAAATGTTTGCCATTG                        2080
B2      ??????????????????????????????????????????????????????????????????????????????????????????????????    2080
SIT1    CCTTACAAGTGTATCTGCTGAATGCCACATAAAAACTAAAAAACTAAATTCATTTACAAACATGTAAATGTTTGCCATTG                        2063
BB1     CCTTACAAGTGTATCTGCTGAATGCCACATAAAAACTAAAAAACTAAATTCATTTACAAACATGTAAATGTTTGCCATTG                        2078
BB2     ??????????????????????????????????????????????????????????????????????????????????????????????????    2080
IMP     CCTTACAAGTGTATCTGCTGAACGCCACATAAAAACTAAAAAACTAAATTCATCTACAAACATGTAAATGTTTGCCATTG                        2071
COW     CCTTACAAGTGTATCTGCTGAATGCCACATAAAAACTCAAAAACTAAATTCATTTACAAATATGTAAATGTTTGCCATTG                        2080
BUF     CCTTAAAAGTGTATCTGCTGAATGCCACATAAAAACTAAAAAACTAAATTCATTTACAAACATGTAAATGTTTGCCATTG                        2079
NIL     CCTTACAAGTGTATCTGCTGAATGCCACATAAAAACTGAAAAACTAAATTCATTTACAAACATGTAAATGTTTGCCATTG                        2080

                2090       2100       2110       2120       2130       2140       2150       2160
                   .          .          .          .          .          .          .          .
GE1     ??????????????????????????????????????????????????????????????????????????????????????????????????    2160
CE1     AGTCAACCTGAATATTCTTCTCATCCTGCTGTTTGGCGGTATCCTTTTCATTAGTAAATGAGGG-ACCAGGAGACAGCCT                        2159
CE2     ??????????????????????????????????????????????????????????????????????????????????????????????????    2160
GK1     AGTCAACCTGAATATTCTTCTCATCCTGCTGTTTGGCGGTATCCTTTTCATTAGTAAATGAGGG-ACCAGGAGACAGCCT                        2158
GK2     ??????????????????????????????????????????????????????????????????????????????????????????????????    2160
LK1     AGTCAACCTGAATATTCTTCTCATCCTGCTGTTTGGCGTTATCCTTTTCATTGGCAAATGAGGG-ACCAGGAGACAGCCT                        2159
CN1     AGTCAACCTGAATATTCTTCTCATCCTGCTGTTTGGCGTTATCCTTTTCATTGGCAAATGAGGG-ACCAGGAGACAGCCT                        2159
CN2     ??????????????????????????????????????????????????????????????????????????????????????????????????    2160
MN1     ??????????????????????????????????????????????????????????????????????????????????????????????????    2160
MN2     ??????????????????????????????????????????????????????????????????????????????????????????????????    2160
B1      AGTCAACCTGAATATTCTTCTCATCCTGCTGTTTGGCGGTATCCTTTTCATTGGCAAATGAGGG-ACCAGGAGACAGCCT                        2159
B2      ??????????????????????????????????????????????????????????????????????????????????????????????????    2160
SIT1    AGTCAACCTGAA-------------------------GGTATCCTTTTCATTGGCAAATGAGGG-ACCAGGAGACAGCCT                        2117
BB1     AGTNAACCTGAATATTCTTCTCATCCTGCTGTTTGGCGGTATCCTTTTCATTGGCAAATTAGGG-ACCAGGAGACAGCCT                        2157
BB2     ??????????????????????????????????????????????????????????????????????????????????????????????????    2160
IMP     AATCAACCTGAATATTCTTCTCATCCTGCTGTTTGGCGATATCCTTTTCATTGGCAAATGAGGG-GCCAGGAGA--GCCT                        2148
COW     AGTCAACCTGAATATTCTTCTCATCCTGCTGTTTGGCGATATCCTTTTCATTGGCAAATGAGGG-ACCAGGAGACAGCCT                        2159
BUF     AGTCAACCTGAATATTCTTCTCATYCTGCTGTTTGGCGGTATCCTTTTCATTGGCAAATGAGGG-ACCAGGAGACAGACT                        2158
NIL     AGTCAACCTGAATATTCTTCTCATCCTGCTGTTTGGCGGTATCCTTTTCATTGGCAAATGAGGGGACCAGGAGACAGCCT                        2160
```

```
                2170       2180       2190       2200       2210       2220       2230       2240
                   .          .          .          .          .          .          .          .
GE1     ????????????????????????????????????????????????????????????????????????????????????????????????    2240
CE1     TTGGGCACTTAGGCCTCTGGTTCTGTTTTCTGGGAGCTGTATTTCGGTCTCAGGAGGGCCCCAGGGGCAGTCTGGGTCAG                    2239
CE2     ????????????????????????????????????????????????????????????????????????????????????????????????    2240
GK1     TTGGGCACTTAGGCCTCTGGTTCTGTTTTCTGGGAGCTGTATTTCGGTCTCAGGAGGGCCCCAGGGGCAGTCTGGGTCAG                    2238
GK2     ????????????????????????????????????????????????????????????????????????????????????????????????    2240
LK1     TTGGGCACTTAGGCCTCTGGTTCTGTTTTCTGGGAGCTGTATTTCGGTCTCAGGAGGGCCCCAGGGGCAGTCTGGGTCAG                    2239
CN1     TTGGGCACTTAGGCCTCTGGTTCTGTTTTCTGGGAGCTGTATTTCGGTCTCAGGAGGGCCCCAGGGGCAGTCTGGGTCAG                    2239
CN2     ????????????????????????????????????????????????????????????????????????????????????????????????    2240
MN1     ????????????????????????????????????????????????????????????????????????????????????????????????    2240
MN2     ????????????????????????????????????????????????????????????????????????????????????????????????    2240
B1      TTGGGCACTTAGGCCCCTGGTTCTGTTTTCTGGTAGCTGTATTTCGGTCTCAGGAGGGCCCCAGGGGCAGTCTGGGTCAG                    2239
B2      ????????????????????????????????????????????????????????????????????????????????????????????????    2240
SIT1    TTGGGCACTTAGGCCTCTGGTTCCGTTTTCTGGGAGCTGTATTTCGGTCTCAGGAGGGCCCCAGGGGCAGTCTGGGTCAG                    2197
BB1     TTGGGAACTTAGGCCTCTGGTTCTGTTTTCTGGGAGCTGTATTTCGGTCTCAGGAGGGCCCCAGGGGCAGTCTGGGTCAG                    2237
BB2     ????????????????????????????????????????????????????????????????????????????????????????????????    2240
IMP     CTGGGCACTTAGGCCTCTGGTTCTGTTTTCTGGGAGCCGTACTGCGGTCTCAAG-GGGCCCCAGGGGCAGTCTGGGTCAG                    2227
COW     TTGGGCACTTAGGCCTCTGGTTCTGTTTTCTGGGAGCTGTATTGCAGTCTCAGGAGGACCCCAGGGGCAGTCTGGGTCAG                    2239
BUF     TTGGGCACTTAGGCCTCTGGTTCTGTTTTCTGGGAGCTGTACTGCGGTCTCAGGAGGACCCCAGGGGCAGTCTGGGTCAG                    2238
NIL     TTGGGCACTTAGGCCTCTGGTTCTGTTTTCTGGGAGCTGTATTGCGGTCTCAGGAGGGCCCCAGGGCAAGTCTGGGTCAG                    2240

                2250       2260       2270       2280       2290       2300       2310       2320
                   .          .          .          .          .          .          .          .
GE1     ????????????????????????????????????????????????????????????????????????????????????????????????    2320
CE1     AC----TCTGGGCAGCATATAGCCTCTGCCAGCTGGACCAGGCTGCCGTAGACCCCCGGCCAGGCAGCGGGCCCTCT--T                    2313
CE2     ????????????????????????????????????????????????????????????????????????????????????????????????    2320
GK1     AC----TCTGGGCAGCATATAGCCTCTGCCAGCTGGACCAGGCTGCCGTAGACCCCCGGCCAGGCAGCGGGCCCTCT--T                    2312
GK2     ????????????????????????????????????????????????????????????????????????????????????????????????    2320
LK1     AC----TCTGGGCAGCATATAGCCTCTGCCAGCTGAACCAGGCTGCCGTAGACCCCCGGCCAGGCAGCGGGCCCTCT--T                    2313
CN1     AC----TCTGGGCAGCATATAGCCTCTGCCAGCTGGACCAGGCTGTC-TAGACCCCCGGCCAGGCAGCGGGCCCTCT--T                    2312
CN2     ????????????????????????????????????????????????????????????????????????????????????????????????    2320
MN1     ????????????????????????????????????????????????????????????????????????????????????????????????    2320
MN2     ????????????????????????????????????????????????????????????????????????????????????????????????    2320
B1      ACAGACTCTGGGCAGCATATAGCCTCTGCCAGCTGGACCAGGCTGCCGTAGACCCCCGGCCAGGCAGCAGGCCCTCT--T                    2317
B2      ????????????????????????????????????????????????????????????????????????????????????????????????    2320
SIT1    AC----TCTGGGCAGCATATAGCCTCTGCCAGCTGGACCAGGCTGCCGTAAACCC--GGCCAGGCAGCGGGCCCTCT--T                    2269
BB1     AT----TCTGGGCAGCATATAGCCTCTGCCAGCTGGACCAGGCTGCCATAGACCCCCGGCCAGGCAGCGGGCCCTCT--T                    2311
BB2     ????????????????????????????????????????????????????????????????????????????????????????????????    2320
IMP     AC----TCTGGGCAGCTCATTGCCTCTGCCAGCTGGACCAGACTGCTG-GGACCCCGGGCCAGGCAGCGGGCTG-CT--T                    2299
COW     AC----TCTGGGCAGC-------CTCTGCCAGCTGGACCAGGCTGCCGTGGACCCCGGGCCAGGCAGCGGGCCCTCT--T                    2306
BUF     AC----TCTGGGCAGC-------CTCTGCCAGCTGGACCAGGCTGCCGTGGACCCCGGGCCAGGCAGCGGGCCCTCT--T                    2305
NIL     AC----TCTGGGCAGCATATTGCCTCTACCAGCTGGACCAGGCTGCCGTGGACCCCGGGCCAAGCAGCGGGCCCTCTCTT                    2316

                2330       2340       2350       2360       2370       2380       2390       2400
                   .          .          .          .          .          .          .          .
GE1     ????????????????????????????????????????????????????????????????????????????????????????????????    2400
CE1     TCAAAACTCCAGGCTGGTTGCTGCGTACAGATGCAAGGGTCTCCGTCTGCCTTAATTGATTCCCCAGCACTTTCGATACC                    2393
CE2     ????????????????????????????????????????????????????????????????????????????????????????????????    2400
GK1     TCAAAACTCCAGGCTGGTTGCTGCATACAGATGCAAGGGTCTCCGTCTGCCTTAACTGATTCCCCAGCACTTTCGATACC                    2392
GK2     ????????????????????????????????????????????????????????????????????????????????????????????????    2400
LK1     TCAAAACTCCAGGCTGGTTGCTGCCTACAGATGCAAGGGTCTCCGTCTGCCTTAACTGATTCCTCAGCACGTTGGATACC                    2393
CN1     TCAAAACTCCAGGCTGGTTGCTGCGTACAGATGCAAGGGTCTCCGTCTGCCTTAACTGATTCCCCAGCACGTTCGATACC                    2392
CN2     ????????????????????????????????????????????????????????????????????????????????????????????????    2400
MN1     ????????????????????????????????????????????????????????????????????????????????????????????????    2400
MN2     ????????????????????????????????????????????????????????????????????????????????????????????????    2400
B1      TCAAAACTCCAGGCTGGTTGCTGCATACAGATGCAAGGGTCTCCGTCTGCCTTAACTGATTCCCCAGCACTTTCGATACC                    2397
B2      ????????????????????????????????????????????????????????????????????????????????????????????????    2400
SIT1    TCAAAACTCCAGGCTGGTTGCTGCATACAGATGCAAGGGTCTCCGTCTGCCTTAACTGATTCCCCAGCACTTTCGATACC                    2349
BB1     TCAAAACTCCAGGCTGGTTGCTGCATACAGATGCAAGGGTCTCGGTCTGCCTTAACGGATTCCCCAGCACTTTCGATACC                    2391
BB2     ????????????????????????????????????????????????????????????????????????????????????????????????    2400
IMP     TCAAAACTCCAGGCCGGCTGCTGCGTGCAGAGGCAAGGGTCTCCGTCTGCCTTAGCTGGTTCCCGAGCACTTTAGATACC                    2379
COW     TCAAAACTCCAGGCTGGCTGCTGCGTGCAGATGCAAGGGTCTCCGTCTGTCTTAACTGGTTCCCAAGCACTTTAGATACC                    2386
BUF     TCAAAACTCCAGGCTGGCTGCTGCGTGCAGATGCAAGGGTCTCCGTCTGCCTTAACTGGTTCCCAAGCACTTTAGATACC                    2385
NIL     TCAAAACTCCAGGCTGGCTGCTGCGTGCAGATGCAAGGGTCTCCGTCTGCCTTAACTGGTTCCCCAGCACTTTAGATACC                    2396
```

```
              2410       2420       2430       2440       2450       2460       2480
                .          .          .          .          .          .          .
GE1     ??????????????????????????????????????????????????????????????CATTTGTTCTCTAAATAAGGAC    2480
CE1     ATCTATACAGTCAAGCTGATCCGAAAAGATTCACCTTAGGACCCCTGCTCTGGGTCCCCATTTGTTCTCTAAATAAGGAC    2473
CE2     ??????????????????????????????????????????????????????????????CATTTGTTCTCTAAATAAGGAC    2480
GK1     ATCTATACAGTCAAGCTGATCCGAAAAGATTCACCTTAGGACCCCTGCTCTGGGTCCCCATTTGTTCTCTAAATAAGGAC    2472
GK2     ??????????????????????????????????????????????????????????????CATTTGTTCTCTAAATAAGGAC    2480
LK1     ATCTATACAGTCAAGCTGATCCGAAAAGATTCACCTTAGGACCCCTGCTCTGGGTCCCCATTTGTTCTCTAAATAAGGAC    2473
CN1     ATCTATACAGTCAAGCTGATCCGAAAAGATTCACCTTAGGACCCCTGCTCTGGGTCCCCATTTGTTCTCTAAATAAGGAC    2472
CN2     ??????????????????????????????????????????????????????????????CATTTGTTCTCTAAATAAGGAC    2480
MN1     ??????????????????????????????????????????????????????????????CATTTGTTCTCTAAATAAGGAC    2480
MN2     ??????????????????????????????????????????????????????????????CATTTGTTCTCTAAATAAGGAC    2480
B1      ATCTATACAGTAATGCTGATCCGAAA------------------------------CATTTGTTCTCTAAATAGGGAC    2445
B2      ??????????????????????????????????????????????????????????????CATTTGTTCTCTAAATAGGGAC    2480
SIT1    ATCTATACAGTCAAGCTGATCCGAAAAGATTCACCTTAGGACCCCTGCTCTGGGTCCCCATTTGTTCTCTAAATAAGGAC    2429
BB1     ATCTATACAGTCAAGCTGATCCGAAAAGATTCACCTTAGGACCCCTGCTCTGGGTCCCCATTTGTTCTCTAAATAAGGAC    2471
BB2     ??????????????????????????????????????????????????????????????CATTTGTTCTCTAAATAAGGAC    2480
IMP     ATCTGTACAGTCAAGACGATCCACAAAGATCCACCTTAGGGCCCCTGCTCTGGGTCCCCATTTGTTCTCTAAATAAGGAC    2459
COW     TTCTCTATAGTCAAGCTGATCCGCAAAGATTCACCCTAGGACCCCTGCTCTGGATCCCCATTTGTTCTCTAAATAAGGAC    2466
BUF     TTCTCTACAGTCAAGCTGATCCGCAAAGATTCACCTTAGGACCCCTGCTCTGGGTCCCCACTTGTTCTCTAAATAAGGAC    2465
NIL     ATCTGTACAGTCAAGCTGATCCGCCAAGATTCACCTTAGGACCCCTGCTCTGGGTCCCCATTTGTTCTCTAAATAGGGAC    2476

              2490       2500       2510       2520       2530       2540       2550       2560
                .          .          .          .          .          .          .          .
GE1     TTGTATGAATGGCCACACGAGGGTTTTACTGTCTCTTACTTCCAATCAGTGAAATTGACCTCCCCGTGAAGAGGCGGGAA    2560
CE1     TTGTATGAATGGCCACACGAGGGTTTTACTGTCTCTTACTTCCAATCAGTGAAATTGACCTCCCCGTGAAGAGGCGGGGA    2553
CE2     TTGTATGAATGGCCACACGAGGGTTTTACTGTCTCTTACTTCCAATCAGTGAAATTGACCTCCCCGTGAAGAGGCGGGGA    2560
GK1     TTGTATGAATGGCCACACGAGGGTTTCACTGTCTCTTACTTCCAATCAGTGAAATTGACCTCCCCGTGAAGAGGCGGGGA    2552
GK2     TTGTATGAATGGCCACACGAGGGTTTCACTGTCTCTTACTTCCAATCAGTGAAATTGACCTCCCCGTGAAGAGGCGGGGA    2560
LK1     TTGTATGAATGGCCACACGAGGGTTTTACTGTCTCTTACTTCCAATCAGTGAAATTGACTTTCCCGTGAAGAGGCGGGAA    2553
CN1     TTGTATGAATGGCCACACGAGGGTTTTACTGTCTCTTACTTCCAATCAGTGAAATTGACCTCCCCGTGAAGAGGCGGGGA    2552
CN2     TTGTATGAATGGCCACACGAGGGTTTTACTGTCTCTTACTTCCAATCAGTGAAATTGACCTCCCCGTGAAGAGGCGGGAA    2560
MN1     TTGTATGAATGGCCACACGAGGGTTCTACTGTCTCTTACTTCCAATCAGTGAAATTGACCTCCCCGTGAAGAGGCGGGAA    2560
MN2     TTGTATGAATGGCCACACGAGGGTTCTACTGTCTCTTACTTCCAATCAGTGAAATTGACCTCCCCGTGAAGAGGCGGGAA    2560
B1      TTGTATGAATGGCCACACGAGGGTTTTACTGTCTCTTACTTCCAATCAGTGAAATTGACCTCCCCGTGAAGAGGCGGGGA    2525
B2      TTGTATGAATGGCCACACGAGGGTTTTACTGTCTCTTACTTCCAATCAGTGAAATTGACCTCCCCGTGAAGAGGCGGGGA    2560
SIT1    TTGTATGAATGGCCACACGAGGGTTTTACTGTCTCTTACTTCCAATCAGTGAAATTGACCTCCCCGTGAAGAGGCGGGGA    2509
BB1     TTGTATGAATGGCCACACGAGGGTTTTACTGTCTCTTACTTCCAATCAGTGAAATTGACCTCCCCGTGAAGAGGCGGGGA    2551
BB2     TTGTATGAATGGCCACACGAGGGTTTTACTGTCTCTTACTTCCAATCAGTGAAATTGACCTCCCCGTGAAGAGGCGGGGA    2560
IMP     TTGTATGAATGGCCACACGAGGGTTTTACTGTCTCTTACTTCCAATCAGTGAAATTGACCTTCCCGTGAAGAGGCGGGAA    2539
COW     TTGTATGAATGGCCGCACGAGGGTTTTACTGTCTCTTACTTCCAATCAGTGAAATTGACCTTCCCGTGAAGAGGCGGGAA    2546
BUF     TTGTATGAATGGCCACACGAGGGTTTCACTGTCTCTTACTTCCAATCAGTGAAATTGACCTTCCCGTGAAGAGGCGGGAA    2545
NIL     TTGTATGAATGGCCACACGAGGGTTTTACTGTCTCTTACTTCCAATCAGTGAAATTGACCTTCCCGTGAAGAGGCGGGGA    2556

              2570       2580       2590       2600       2610       2620       2630       2640
                .          .          .          .          .          .          .          .
GE1     TAAATTAATAAGACGAGAAGACCCTATGGAGCTTTAACTAACCAACTCAAAGAAAA-CACAATAA-ATCACTCTGGGAAT    2638
CE1     TAAATTAATAAGACGAGAAGACCCTATGGAGCTTTAACTAACCAATTCAAAGAAAA-CACAATAA-ATCACCCTGGGAAT    2631
CE2     TAAATTAATAAGACGAGAAGACCCTATGGAGCTTTAACTAACCAATTCAAAGAAAA-CACAATAA-ATCACCCTGGGAAT    2638
GK1     TAAACTAACAAGACGAGAAGACCCTATGGAGCTTTAACTAACCAATTTAAAGAAAA-CATAATAA-ACCACTCAGGGAAT    2630
GK2     TAAACTAACAAGACGAGAAGACCCTATGGAGCTTTAACTAACCAATTTAAAGAAAA-CATAATAA-ACCACTCAGGGAAT    2638
LK1     TAAACTAATAAGACGAGAAGACCCTATGGAGCTTCAACTAATCAGTCCAAAGAAAA-TAAACTAATATTGCTCAGGAAAT    2632
CN1     TAAACTAATAAGACGAGAAGACCCTATGGAGCTTTAACTAACCAATTCAAAGAAAA-CATAATAA-ATCACCCAAGAGAT    2630
CN2     TAAACTAATAAGACGAGAAGACCCTATGGAGCTTTAACTAACCAATTCAAAGAAAA-CATAATAA-ATCACCCAGGAAAT    2638
MN1     TGAATAAATAAGACGAGAAGACCCTATGGAGCTTTAACTAATCAACTCAAAGAAAA-CATAATAA-ATCACTCAGGGAAT    2638
MN2     TGAATAAATAAGACGAGAAGACCCTATGGAGCTTTAACTAATCAACTCAAAGAAAA-CATAATAA-ATCACTCAGGGAAT    2638
B1      TGAACTAATAAGACGAGAAGACCCTATGGAGCTTTAACTAACCAATTCAAAGAAAG-CATAATAA-ACCACCTAGGAAAT    2603
B2      TGAACTAATAAGACGAGAAGACCCTATGGAGCTTTAACTAACCAATTCAAAGAAAA-CATAATAA-ACCACCTAGGAAAT    2638
SIT1    TGAACCAATAAGACGAGAAGACCCTATGGAGCTTTAACTAACCAATTCAAAGAAAA-CATAATAA-ATCACTCAGGAAAT    2587
BB1     TGAACTAATAAGACGAGAAGACCCTATGGAGCTTTAACTAACCAATTCAAAGAAAA-TATAAAAA-ATCACCCAAGAGAT    2629
BB2     TAAACTAATAAGACGAGAAGACCCTATGGAGCTTTAACTAACCAATTCAAAGAAAA-TATAAAAA-ATCACCCAAGAGAT    2638
IMP     TAAATAAATAAGACGAGAAGACCCTATGGAGCTTTAACTAACTAGTCCAAAGAGAA-TAAACCAA-ACCACC-AAGGAAT    2616
COW     TGCACAAATAAGACGAGAAGACCCTATGGAGCTTTAACTAACCAACCCAAAGAGAA-TAGATTTA-ACCATT-AAGGAAT    2623
BUF     TATACTAATAAGACGAGAAGACCCTATGGAGCTTTAACTAATCAGCCCAAAGAAAAATAAACCAA-ACCAC-CAAGGAAT    2623
NIL     TGAACCAATAAGACGAGAAGACCCTATGGAGCTTCAACTAACTAGCCCAAAGAAAA-CA-AATACAACCACCCAAGGAAT    2634
```

```
                2650       2660       2670       2680       2690       2700       2710       2720
                 .          .          .          .          .          .          .          .
GE1     AGCAATGCTTTCTTATGAAT-TGGCAGTTTCGGTTGGGGTGACCTCGGAGAACAAAAAATCCTCCGAACGATTTTAAAGA    2717
CE1     AACAATATTTTCTCATGAAT-TGGCAGTTTCGGTTGGGGTGACCTCGGAGAATAAAAAATCCTCCGAACGATTTTAAAGA    2710
CE2     AACAATATTTTCTCATGAAT-TGGCAGTTTCGGTTGGGGTGACCTCGGAGAATAAAAAATCCTCCGAACGATTTTAAAGA    2717
GK1     AACAATATTTTCTTATAAGT-TGGCAGTTTCGGTTGGGGTGACCTCGGAGAATAAAAAATCCTCCGAGCGATTTTAAAGA    2709
GK2     AACAATATTTTCTTATAAGT-TGGCAGTTTCGGTTGGGGCGACCTCGGAGAATAAAAAATCCTCCGAGCGATTTTAAAGA    2717
LK1     AACAATATTTTCCTATGGCC-TGACAGTTTCGGTTGGGGCGACCTCGGAGAATAAAAAATCCTCCGAACGATTATAAAGA    2711
CN1     AACAACATTTTCTTATGAGT-TGACAGTTTCGGTTGGGGTGACCTCGGAGAATAAAAAATCCTCCGAACGATTCTAAAGA    2709
CN2     AACAATATTTTCTTATGAGT-TGGCAGTTTCGGTTGGGGTGACCTCGGAGAATAAAAAATCCTCCGAACGATTCTAAAGA    2717
MN1     AACAATATTTTCTTATGAAT-TGACAGTTTCGGTTGGGGTGACCTCGGAGAATAAAAAATCCTCCGAACGATTCTAAAGA    2717
MN2     AACAATATTTTCTTATGAAT-TGACAGTTTCGGTTGGGGTGACCTCGGAGAATAAAAAATCCTCCGAACGATTCTAAAGA    2717
B1      AACAACATTTTCTCATGAGT-TGGCAGTTTCGGTTGGGGTGACCTCGGAGAATAAAAAATCCTCCGAACGATTTTAAAGG    2682
B2      AACAACATTTTCTCATGAGT-TGGCAGTTTCGGTTGGGGTGACCTCGGAGAATAAAAAATCCTCCGAACGATTTTAAAGG    2717
SIT1    AACAATATTTTCTCATGAAT-TGGCAGTTTCGGTTGGGGTGACCTCGGAGAATAAAAAATCCTCCGAACGATTTTAAAGA    2666
BB1     AACAACATTTTCTCATGAGT-TGACAGTTTCGGTTGGGGTGACCTCGGAGAATAAAAAATCCTCCGAACGATTCTAAAGA    2708
BB2     AACAACATTTTCTCATGAGT-TGACAGTTTCGGTTGGGGTGACCTCGGAGAATAAAAAATCCTCCGAACGATTCTAAAGA    2717
IMP     AAAAATACTCCTG-ATGGAC-TAGCAGTTTTGGTTGGGGTGACCTCGGAGAACAAAAAATCCTCCGAACGATTCTAAAAA    2694
COW     AACAACAATCTCC-ATGAGT-TGGTAGTTTCGGTTGGGGTGACCTCGGAGAATAAAAAATCCTCCGAGCGATTTTAAAGA    2701
BUF     AACAATACCCTCC-ATGGGC-TAACAGTTTCGGTTGGGGTGACCTCGGAGAATAAAAAATCCTCCGAGCGATTTTAAAGA    2701
NIL     A-AAAAACTCTCT-ATGGGCATAGCAGTTTCGGTTGGGGTGACCTCGGAGAACAAAAAATCCTCCGAGCGATTTTAAAGA    2712

                2730       2740       2750       2760       2770       2780       2790       2800
                 .          .          .          .          .          .          .          .
GE1     CTAGACCTACAAGTCAACTCACTCAATCGCTTATTGATCCAAAAAGCTTGATCAACGGAACAAGTTACCCTAGGGATAAC    2797
CE1     CTAGACCTACAAGTCAACCCACTCAATCGCTTATTGATCCAAAAAGCTTGATCAACGGAACAAGTTACCCTAGGGATAAC    2790
CE2     CTAGACCTACAAGTCAACCCACTCAATCGCTTATTGATCCAAAAAGCTTGATCAACGGAACAAGTTACCCTAGGGATAAC    2797
GK1     CTAGACCTACAAGTCAACTCACTTAATCGCTTATTGATCCAAAAAGCTTGATCAACGGAACAAGTTACCCTAGGGATAAC    2789
GK2     CTAGACCTACAAGTCAACTCACTTAATCGCTTATTGATCCAAAAAGCTTGATCAACGGAACAAGTTACCCTAGGGATAAC    2797
LK1     CTAGACCTACAAGTCAAATCACTCAATCGCTTATTGATCCAAAAA-CTTGATCAACGGAACAAGTTACCCTAGGGATAAC    2790
CN1     CTAGACCCACAAGTCAAATCACTCAATCGCTCATTGATCCAAAAAACTTGATCAACGGAACAAGTTACCCTAGGGATAAC    2789
CN2     CTAGACCTACAAGTCAAATCACTCAATCGCTCATTGATCCAAAAAACTTGATCAACGGAACAAGTTACCCTAGGGATAAC    2797
MN1     CTAGACCTACAAGTCAAATCACTCAATCGCTTATTGATCCAAAAAGCTTGATCAACGGAACAAGTTACCCTAGGGATAAC    2797
MN2     CTAGACCTACAAGTCAAATCACTCAATCGCTTATTGATCCAAAAAGCTTGATCAACGGAACAAGTTACCCTAGGGATAAC    2797
B1      CTAGACCTACAAGTCAAATCACTCAGTCGCTTATTGATCCAAGAAACTTGATCAACGGAACAAGTTACCCTAGGGATAAC    2762
B2      CTAGACCTACAAGTCAAATCACTCAGTCGCTTATTGATCCAAGAAACTTGATCAACGGAACAAGTTACCCTAGGGATAAC    2797
SIT1    CTAGACCTACAAGTCAAATCACTCAATCGCTTATTGATCCAAAAAACTTGATCAACGGAACAAGTTACCCTAGGGATAAC    2746
BB1     CTAGACCCACAAGTCAAATCACTCAATCGCTCATTGATCCAAAAAATTTGATCAACGGAACAAGTTACCCTAGGGATAAC    2788
BB2     CTAGACCCACAAGTCAAATCACTCAATCGCTCATTGATCCAAAAAGTTTGATCAACGGAACAAGTTACCCTAGGGATAAC    2797
IMP     CCAGACTCACCAGTCAAATTATTCTATCGCTTATTGACCCAAAAA-TTTGATCAACGGAACAAGTTACCCTAGGGATAAC    2773
COW     CTAGACCCACAAGTCAAATCACTCTATCGCTCATTGATCCAAAAA-CTTGATCAACGGAACAAGTTACCCTAGGGATAAC    2780
BUF     TTAGACCTACAAGTCAAATCACTCTATCGCTTATTGATCCAAAAA-CTTGATCAACGGAACAAGTTACCCTAGGGATAAC    2780
NIL     TTAGACCTACAAGCCAAATCACTCTATCGCTTATTGATCCAAAAA-CTTGATCAACGGAATAAGTTACCCTAGGGATAAC    2791

                2810       2820       2830       2840       2850       2860       2870       2880
                 .          .          .          .          .          .          .          .
GE1     AGCGCAAACTTACACATGCAAGCATCCACATCCCAGTGAGAATGCCCTCCAGGTCAATAAGACTAAGAGGAGCTGGTATC    2877
CE1     AGCGCAAACTTACACATGCAAGCATCCTCATCCCAGTGAGAATGCCCTCCAGGTCAATAAGACTAAGAGGAGCTGGTATC    2870
CE2     AGCGCAAACTTACACATGCAAGCATCCTCATCCCAGTGAGAATGCCCTCCAGGTCAATAAGACTAAGAGGAGCTGGTATC    2877
GK1     AGCGCAAACTTACACATGCAAGCATCCACATCCCAGTGAGAATGCCCTCCAGGTCAATAAGACTAAGAGGAGCTGGTATC    2869
GK2     AGCGCAAACTTACACATGCAAGCATCCACATCCCAGTGAGAATGCCCTCCAGGTCAATAAGACTAAGAGGAGCTGGTATC    2877
LK1     AGCGCAAACTTACACATGCAAGCATCCACATCCCAGTGAGAATGCCCTCCAGGTCAACAAGACTAAGAGGAGCTGGTATC    2870
CN1     AGCGCAAACTTACACATGCAAGCATCCACATCCCAGTGAGAATGCCCTCCAAGTCAATAGGACTAAGAGGAGCTGGTATC    2869
CN2     AGCGCAAACTTACACATGCAAGCATCCACATCCCAGTGAGAATGCCCTCCAGGTCAATAGGACTAAGAGGAGCTGGTATC    2877
MN1     AGCGCAAACTTACACATGCAAGCATCCACACCCCAGTGAGAATGCCCTCCGGGTCAATAAGACTAAAAGGAGCTGGTATC    2877
MN2     AGCGCAAACTTACACATGCAAGCATCCACACCCCAGTGAGAATGCCCTCCGGGTCAATAAGACTAAAAGGAGCTGGTATC    2877
B1      AGCGCAAACTTACACATGCAAGCATCCACATCCCAGTGAGAATGCCCTCCAAGTCAACAAGACTAAGAGGAGCTGGTATC    2842
B2      AGCGCAAACTTACACATGCAAGCATCCTCATCCCAGTGAGAATGCCCTCCAAGTCAATAAGACTAAGAGGAGCTGGTATC    2877
SIT1    AGCGCAAACTTACACATGCAAGCATCCACATCCCAGTGAGAATGCCCTCCAGGTCAATAAGACTGAGAGGAGCTGGTATC    2826
BB1     AGCGCAAACTTACACATGCAAGCATCCTCATCCCAGTGAGAATGCCCTCCAGGTCAATAAGACTAAGAGGAGCTGGTATC    2868
BB2     AGCGCAAACTTACACATGCAAGCATCCTCATCCCAGTGAGAATGCCCTCCAGGTCAATAAGACTAAGAGGAGCTGGTATC    2877
IMP     AGCGCAAACTTACACATGCAAGCATCTACACCCCGGTGAGAATGCCCTCTGAATCACCAAGACTAAGAGGAGCGGGTATC    2853
COW     AGCGCAAACTTACACATGCAAGCATCTACACCCCAGTGAGAATGCCCTCTAGGTTATTAAAACTAAGAGGAGCTGGCATC    2860
BUF     AGCGCAAACTTACACATGCAAGCATCCACACCCCGGTGAGAATGCCCTCTAGGTTAATAAAACCAAGAGGAGCGGGTATC    2860
NIL     AGCGCAAACTTACACATGCAAGCATCCACATCCCGGTGAGAATGCCCTCTAGGTCAATAAGACTAAGAGGAGCTGGTATC    2871
```

```
                2890       2900       2910       2920       2930       2940       2950       2960
                   .          .          .          .          .          .          .          .
GE1     AAGCACACACTCGTAGCTCACGACACCTTGCTTAACCACACCCCCACGGGAAACAGCAGTGATAAAAATTAAGCCATAAA    2957
CE1     AAGCACACACTCGTAGCTCACGACACCTTGCTTAACCACACCCCCACGGGAAACAGCAGTGATAAAAATTAAGCCATAAA    2950
CE2     AAGCACACACTCGTAGCTCACGACACCTTGCTTAACCACACCCCCACGGGAAACAGCAGTGATAAAAATTAAGCCATAAA    2957
GK1     AAGCACACACTCGTAGCTCACGACACCTTGCTTAACCACACCCCCACGGGAAACAGCAGTGATAAAAATTAAGCCATAAA    2949
GK2     AAGCACACACTCGTAGCTCACGACACCTTGCTTAACCACACCCCCACGGGATACAGCAGTGATAAAAATTAAGCCATAAA    2957
LK1     AAGCACACACTCGTAGCTCACGACACCTTGCTTAACCACACCCCCACGGGAAACAGCAGTGATAAAAATTAAGCCATAAA    2950
CN1     AAGCACACACTCGTAGCTCACGACACCTTGCTTAACCACACCCCCACGGGAAACAGCAGTGATAAAAATTAAGCCATAAA    2949
CN2     AAGCACACACTCGTAGCTCACGACACCTTGCTTAACCACACCCCCACGGGAAACAGCAGTGATAAAAATTAAGCCATAAA    2957
MN1     AAGCACACACCCGTAGCTCACGACACCTTGCTTAACCACACCCCCACGGGATACAGCAGTGATAAAAATTAAGCCATAAA    2957
MN2     AAGCACACACCCGTAGCTCACGACACCTTGCTTAACCACACCCCCACGGGATACAGCAGTGATAAAAATTAAGCCATAAA    2957
B1      AAGCACACACCCGTAGCTCACGACACCTTGCTTAACCACACCCCCACGGGAAACAGCAGTGATAAAAATTAAGCCATAAA    2922
B2      AAGCACACACCCGTAGCTCACGACACCTTGCTTAACCACACCCCCACGGGATACAGCAGTGATAAAAATTAAGCCATAAA    2957
SIT1    AAGCACACACTCGTAGCTCACGACACCTTGCTTAACCACACCCCCACGGGAAACAGCAGTGATAAAAATTAAGCCATAAA    2906
BB1     AAGCACACACTCGTAGCTCACGACACCTTGCTTAACCACACCCCCACGGGATACAGCAGTGATAAAAATTAAGCCATAAA    2948
BB2     AAGCACACACTCGTAGCTCACGACACCTTGCTTAACCACACCCCCACGGGAAACAGCAGTGATAAAAATTAAGCCATAAA    2957
IMP     AAGCACACACTTGTAGCTCATGACACCTTGCTTAACCACACCCCCACGGGAGACAGCAGTGACAAAAATTAAGCCATAAA    2933
COW     AAGCACACACCCGTAGCTCACGACGCCTTGCTTAACCACACCCC-ACGGGAAACAGCAGTGACAAAAATTAAGCCATAAA    2939
BUF     AAGCACACACCTGTAGCTCACGACGCCTTGCTTAACCACACCCCCACGGGAGACAGCAGTGACAAAAATTAAGCCATAAA    2940
NIL     AAGCACACATCCGTAGCTCACGACACCTTGCTTAACCACACCCCCACGGGAAACAGCAGTGATAAAAATTAAGCTATAAA    2951

                2970       2980       2990       3000       3010       3020       3030       3040
                   .          .          .          .          .          .          .          .
GE1     CGAAAGTTTGACTAAGTTATGTTAATTAGGGTTGGTAAATTTCGTGCCAGCCACCGCGGCCATACGATTAACCCAAGCTA    3037
CE1     CGAAAGTTTGACTAAGTTATGTTAATTAGGGTTGGTAAATTTCGTGCCAGCCACCGCGGCCATACGATTAACCCAAGCTA    3030
CE2     CGAAAGTTTGACTAAGTTATGTTAATTAGGGTTGGTAAATTTCGTGCCAGCCACCGCGGCCATACGATTAACCCAAGCTA    3037
GK1     CGAAAGTTTGACTAAGTTATGTTAATTAGGGTTGGTAAATTTCGTGCCAGCCACCGCGGCCATACGATTAACCCAAGCTA    3029
GK2     CGAAAGTTTGACTAAGTTATGTTAATTAGGGTTGGTAAATTTCGTGCCAGCCACCGCGGCCATACGATTAACCCAAGCTA    3037
LK1     CGAAAGTTTGACTAAGTTATGTTAATTAGGGTTGGTAAATTTCGTGCCAGCCACCGCGGCCATACGATTAACCCAAGCTA    3030
CN1     CGAAAGTTTGACTAAGTTATGTTAATTAGGGTTGGTAAATTTCGTGCCAGCCACCGCGGCCATACGATTAACCCAAGCTA    3029
CN2     CGAAAGTTTGACTAAGTTATGTTAATTAGGGTTGGTAAATTTCGTGCCAGCCACCGCGGCCATACGATTAACCCAAGCTA    3037
MN1     CGAAAGTTTGACTAAGTTATGTTAATTAGGGTTGGTAAATTTCGTGCCAGCCACCGCGGCCATACGATTAACCCAAGCTA    3037
MN2     CGAAAGTTTGACTAAGTTATGTTAATTAGGGTTGGTAAATTTCGTGCCAGCCACCGCGGCCATACGATTAACCCAAGCTA    3037
B1      CGAAAGTTTGACTAAGTTATGTTAATTAGGGTTGGTAAATTTCGTGCCAGCCACCGCGGCCATACGATTAACCCAAGCTA    3002
B2      CGAAAGTTTGACTAAGTTATGTTAATTAGGGTTGGTAAATTTCGTGCCAGCCACCGCGGCCATACGATTAACCCAAGCTA    3037
SIT1    CGAAAGTTTGACTAAGTTATGTTAATTAGGGTTGGTAAATTTCGTGCCAGCCACCGCGGCCATACGATTAACCCAAGCTA    2986
BB1     CGAAAGTTTGACTAAGTTATGTTAATTAGGGTTGGTAAATTTCGTGCCAGCCACCGCGGCCATACGATTAACCCAAGCTA    3028
BB2     CGAAAGTTTGACTAAGTTATGTTAATTAGGGTTGGTAAATTTCGTGCCAGCCACCGCGGCCATACGATTAACCCAAGCTA    3037
IMP     CGAAAGTTTGACTAAGTTATATTGACTAGGGTTGGTAAATTTCGTGCCAGCCACCGCGGTCATACGATTAACCCAAGTTA    3013
COW     CGAAAGTTTGACTAAGTTATATTAATTAGGGTTGGTAAATCTCGTGCCAGCCACCGCGGTCATACGATTAACCCAAGCTA    3019
BUF     CGAAAGTTTGACTAAGTTATGTTAATAAGGGTTGGTAAATCTCGTGCCAGCCACCGCGGTCATACGATTAACCCAAGCTA    3020
NIL     CGAAAGTTTGACTAAGTTATATTAATTAGGGTTGGTAAATTTCGTGCCAGCCACCGCGGCCATACGATTAACCCAAGCTA    3031

                3050       3060       3070       3080       3090       3100       3110       3120
                   .          .          .          .          .          .          .          .
GE1     ACAGGAATACGGCGTAAAACGTGTTTAAGCACCACACCAAATAGAGTTAAATTTTAATTAAGCTGTAAAAAGCCATAATT    3117
CE1     ACAGGAATACGGCGTAAAACGTGTTTAAGCACCACACCAAATAGAGTTAAATTTTAATTAAGCTGTAAAAAGCCATAATT    3110
CE2     ACAGGAATACGGCGTAAAACGTGTTTAAGCACCACACCAAATAGAGTTAAATTTTAATTAAGCTGTAAAAAGCCATAATT    3117
GK1     ACAGGAATACGGCGTAAAACGTGTTTAAGCGCCACACCAAATAGAGTTAAATTTTAATTAAGCTGTAAAAAGCCGTAATT    3109
GK2     ACAGGAATACGGCGTAAAACGTGTTTAAGCGCCACACCAAATAGAGTTAAATTTTAATTAAGCTGTAAAAAGCCGTAATT    3117
LK1     ACAGGAATACGGCGTAAAACGTGTTTAAGCACCACACCAAATAGAGTTAAATTTTAATTAAGCTGTAAAAAGCCATAATT    3110
CN1     ACAGGAATACGGCGTAAAACGTGTTTAAGCACCACACCAAATAGAGTTAAATCTTAATTAAGCTGTAAAAAGCCATAATT    3109
CN2     ACAGGAATACGGCGTAAAACGTGTTTAAGCACCACACCAAATAGAGTTAAATTTTAATTAAGCTGTAAAAAGCCATAATT    3117
MN1     ACGGGAGTACGGCGTAAAACGTGTTCAAGCACTACACCAAATAGAGTTAAATCCTAGTTAAGCTGTAAAAAGCCATAATT    3117
MN2     ACGGGAGTACGGCGTAAAACGTGTTCAAGCACTACACCAAATAGAGTTAAATCCTAGTTAAGCTGTAAAAAGCCATAATT    3117
B1      ACAGGAATACGGCGTAAAACGTGTTTAAGCGCCACACCAAATAGAGTTAAATTTTAATTAAGCTGTAAAAAGCCATAATT    3082
B2      ACGGGAATACGGCGTAAAAAGTGTTTAAGCGCCACACCAAATAGAGTTAAACTTTAATTAAGCTGTAAAAAGCCATAATT    3117
SIT1    ACAGGAATACGGCGTAAAACGTGTTTAAGCACCACACCAAATAGAGTTAAATCTTAATTAAGCTGTAAAAAGCCATAATT    3066
BB1     ACAGGAATACGGCGTAAAACGTGTTTAAGCGCCACACCAAATAGAGTTAAATTTTAATTAAGCTGTAAAAAGCCATAATT    3108
BB2     ACAGGAATACGGCGTAAAACGTGTTTAAGCACCACACCAAATAGAGTTAAATTTTAATTAAGCTGTAAAAAGCCATAATT    3117
IMP     ACGGGAATACGGCGTAAAACGTGTTTAAGCATTACACTAAATAGAGTTAAATCCTAATTAAACTGTAAAAAGCCATAATT    3093
COW     ACAGGAATACGGCGTAAAACGTGTTTAAGCACCATACCAAATAGGGTTAAATTCTAACTAAGCTGTAAAAAGCCATGATT    3099
BUF     ACGGGAACACGGCGTAAAATGTGTTAAAGCATCCTACTAAGTAGAGTTAAATTTTAATTAAGCTGTAAAAAGCCATAATT    3100
NIL     ACAGGAATACGGCGTAAAACGTGTTTAAGCACCACACCAAATAGAGTTAAATTTTAATTAAGCTGTAAAAAGCCATAATT    3111
```

```
                3130       3140       3150       3160       3170       3180       3190       3200
                 .          .          .          .          .          .          .          .
GE1     ATTGTAAAAATAAATAACGAAAGTGACTCTACAACAGCTGACACACCATAGCTAAGACCCAAACTGGGATTAGATACCCC    3197
CE1     ATTGTAAAAATAAATAACGAAAGTGACTCTACAACAGCTGACACACCATAGCTAAGACCCAAACTGGGATTAGATACCCC    3190
CE2     ATTGTAAAAATAAATAACGAAAGTGACTCTACAACAGCTGACACACCATAGCTAAGACCCAAACTGGGATTAGATACCCC    3197
GK1     ATTGTAAAAATAAATAACGAAAGTGACTCTACAACAGCTGACACACCATAGCTAAGACCCAAACTGGGATTAGATACCCC    3189
GK2     ACTGTAAAAATAAATAACGAAAGTGACTCTACAACAGCTGACACACCATAGCTAAGACCCAAACTGGGATTAGATACCCC    3197
LK1     ATTGTAAAAATAAATAACGAAAGTGACTCTACAACAGCTGACACACCATAGCTAAGACCCAAACTGGGATTAGATACCCC    3190
CN1     ATTGTAAAAATAAATAACGAAAGTGACTCTACAACAGCTAACACACCATAGCTAAGACCCAAACTGGGATTAGATACCCC    3189
CN2     ATTGTAAAAATAAATAACGAAAGTGACTCTACAACAGCTGACACACCATAGCTAAGACCCAAACTGGGATTAGATACCCC    3197
MN1     ATTGTAAAAATAAACGACGAAAGTGACTCTACAACAGCTGATACACCATAGCTAAGACCCAAACTGGGATTAGATACCCC    3197
MN2     ATTGTAAAAATAAACGACGAAAGTGACTCTACAACAGCTGATACACCATAGCTAAGACCCAAACTGGGATTAGATACCCC    3197
B1      ATTGTAAAAATAAATAACGAAAGTAACTCTACAACAGCTGACACACCATAGCTAAGACCCAAACTGGGATTAGATACCCC    3162
B2      ATTGTAAAAATAAATAACGAAAGTGACTCTACAGCAGCTGACACACCATAGCTAAGACCCAAACTGGGATTAGATACCCC    3197
SIT1    ATTGTAAAAATAAATAACGAAAGTAACTCTACAACAGCTGACACACCATAGCTAAGACCCAAACTGGGATTAGATACCCC    3146
BB1     ACTGTAAAAATAAATAACGAAAGTGACTCTACAACAGCTGACACACCATAGCTAAGACCCAAACTGGGATTAGATACCCC    3188
BB2     ACTGTAAAAATAAATAACGAAAGTGACTCTACAACAGCTGACACACCATAGCTAAGACCCAAACTGGGATTAGATACCCC    3197
IMP     ATAGTGAAAGTAAACGACGAAAGTAACTCTACTACAGCTGACACACTATAGCTAAGACCCAAACTGGGATTAGATACCCC    3173
COW     AAAATAAAAATAAATGACGAAAGTGACCCTACAACAGCTGACGCACTATAGCTAAGACCCAAACTGGGATTAGATACCCC    3179
BUF     TCAATAAAAATAAATGACGAAAGTAACTCTATAGCAGCTGACACACTATAGCTAAGATCCAAACTGGGATTAGATACCCC    3180
NIL     ACTGTAAAAATAAATAACGAAAGTGACTCTACAACAGCTGACACACCATAGCTAAGACCCAAACTGGGATTAGATACCCC    3191

                3210       3220       3230       3240       3250       3260       3270       3280
                 .          .          .          .          .          .          .          .
GE1     ACTATGCTTAGCCCTAAACACAGATAATTATATAAACAAAATTATTCGCCAGAGTACTACTAGCAATAGCTTAAAACTCA    3277
CE1     ACTATGCTTAGCCCTAAACACAGATAATTATATAAACAAAATTATTCGCCAGAGTACTACTAGCAATAGCTTAAAACTCA    3270
CE2     ACTATGCTTAGCCCTAAACACAGATAATTATATAAACAAAATTATTCGCCAGAGTACTACTAGCAATAGCTTAAAACTCA    3277
GK1     ACTATGCTTAGCCCTAAACACAGATAATTATATAAACAAAATTATTCGCCAGAGTACTACTAGCAATAGCTTAAAACTCA    3269
GK2     ACTATGCTTAGCCCTAAACACAGATAATTATATAAACAAAATTATTCGCCAGAGTACTACTAGCAATAGCTTAAAACTCA    3277
LK1     ACTATGCTTAGCCCTAAACACAGATAATTATATAAACAAAATTATTCGCCAGAGTACTACTAGCAATAGCTTAAAACTCA    3270
CN1     ACTATGCTTAGCCCTAAACACAGATAATTATATAAACAAAATTATTCGCCAGAGTACTACTAGCGATAGCTTAAAACTCA    3269
CN2     ACTATGCTTAGCCCTAAACACAGATAATTATATAAACAAAATTATTCGCCAGAGTACTACTAGCGATAGCTTAAAACTCA    3277
MN1     ACTATGCTTAGCCCTAAACACAAATAATTATATAAACAAAATTATTCGCCAGAGTACTACTAGCAATAGCTTAAAACTCA    3277
MN2     ACTATGCTTAGCCCTAAACACAGATAATTATATAAACAAAATTATTCGCCAGAGTACTACTAGCAATAGCTTAAAACTCA    3277
B1      ACTATGCTTAGCCCTAAACACAGATAATTATATAAACAAAATTATTCGCCAGAGTACTACTAGCAACAGCTTAAAACTCA    3242
B2      ACTATGCTTAGCCCTAAACACAGATAATTATATAAACAAAATTATTCGCCAGAGTACTACTAGCAACAGCTTAAAACTCA    3277
SIT1    ACTATGCTTAGCCCTAAACACAGATAATTATATAAACAAAATTATTCGCCAGAGTACTACTAGCAATAGCTTAAAACTCA    3226
BB1     ACTATGCTTAGCCCTAAACACAGATAATTATATAAACAAAATTATTCGCCAGAGTACTACTAGCAATAGCTTAAAACTCA    3268
BB2     ACTATGCTTAGCCCTAAACACAGATAATTATATAAACAAAATTATTCGCCAGAGTACTACTAGCAATAGCTTAAAACTCA    3277
IMP     ACTATGCTCAGCCCTAAACACAAATAATCATA-AAACAAGATTATTCGCCAGAGTACTACCGGCAACGGCCCAAAACTCA    3252
COW     ACTATGCTTAGCCCTAAACACAGATAATTACATAAACAAAATTATTCGCCAGAGTACTACTAGCAACAGCTTAAAACTCA    3259
BUF     ACTATGCTTAGCCCTAAACACAGATAATTATATTAACAAAATTATTCGCCAGGGTACTACCGGCAATAGCCTAAAACTCA    3260
NIL     ACTATGCTTAGCCCTAAACACAGATAATTATATAAACAAAATTATTCGCCAGAGTACTACTGGCAACAGCTTAAAACTCA    3271

                3290       3300       3310       3320       3330       3340       3350       3360
                 .          .          .          .          .          .          .          .
GE1     AAGGACTTGGCGGTGCTTCATATCCCTCTAGAGGAGCCTGTTCTATAATCGATAAACCCCGATAAACCTCACCAATTCTT    3357
CE1     AAGGACTTGGCGGTGCTTCATACCCCTCTAGAGGAGCCTGTTCTATAATCGATAAACCCCGATAAACCTCACCAATTCTT    3350
CE2     AAGGACTTGGCGGTGCTTCATATCCCTCTAGAGGAGCCTGTTCTATAATCGATAAACCCCGATAAACCTCACCAATTCTT    3357
GK1     AAGGACTTGGCGGTGCTTCATATCCCTCTAGAGGAGCCTGTTCTATAATCGATAAACCCCGATAAACCTCACCAGTTCTT    3349
GK2     AAGGACTTGGCGGTGCTTCATATCCCTCTAGAGGAGCCTGTTCTATAATCGATAAACCCCGATAAACCTCACCAATTCTT    3357
LK1     AAGGACTTGGCGGTGCTTCATATCCCTCTAGAGGAGCCTGTTCTATAATCGATAAACCCCGATAAACCTCACCAATTCTT    3350
CN1     AAGGACTTGGCGGTGCTTCATATCCCTCTAGAGGAGCCTGTTCTATAATCGATAAACCCCGATAAACCTCACCAATTCTT    3349
CN2     AAGGACTTGGCGGTGCTTCATATCCCTCTAGAGGAGCCTGTTCTATAATCGATAAACCCCGATAAACCTCACCAATTCTT    3357
MN1     AAGGACTTGGCGGTGCTTCATATCCCTCTAGAGGAGCCTGTTCTATAATCGATAAACCCCGATAAACCTCACCAATTCTT    3357
MN2     AAGGACTTGGCGGTGCTTCATATCCCTCTAGAGGAGCCTGTTCTATAATCGATAAACCCCGATAAACCTCACCAATTCTT    3357
B1      AAGGACTTGGCGGTGCTTCATATCCCTCTAGAGGAGCCTGTTCTATAATCGATAAACCCCGATAAACCTCACCAATTCTT    3322
B2      AAGGACTTGGCGGTGCTTCATATCCCTCTAGAGGAGCCTGTTTTATAATCGATAAACCCCGATAAACCTCACCAATTCTT    3357
SIT1    AAGGACTTGGCGGTGCTTCATATCCCTCTAGAGGAGCCTGTTCTATAATCGATAAACCCCGATAAACCTCACCAATTCTT    3306
BB1     AAGGACTTGGCGGTGCTTCATATCCCTCTAGAGGAGCCTGTTCTATAATCGATAAACCCCGATAAACCTCACCAATTCTT    3348
BB2     AAGGACTTGGCGGTGCTTCATATCCCTCTAGAGGAGCCTGTTCTATAATCGATAAACCCCGATAAACCTCACCAATTCTT    3357
IMP     AAGGACTTGGCGGTGCTTTATACCCTTCTAGAGGAGCCTGTTCTATAATCGATAAACCCCGATAAACCTCACCATTTCTT    3332
COW     AAGGACTTGGCGGTGCTTTATATCCTTCTAGAGGAGCCTGTTCTATAATCGATAAACCCCGATAAACCTCACCAATTCTT    3339
BUF     AAGGACTTGGCGGTGCTTTATATCCCTCTAGAGGAGCCTGTTCTATAATCGATAAACCCCGATAAACCTCACCAATTCTT    3340
NIL     AAGGACTTGGCGGTGCTTCATATCCCTCTAGAGGAGCCTGTTCTATAATCGATAAACCCCGATAAACCTCACCAATTCTT    3351
```

```
                3370       3380       3390       3400       3410       3420       3430       3440
                 .          .          .          .          .          .          .          .
GE1     GCTAATACAGTCTATATACCGCCATCTTCAGCAAACCATTTATGCAATGAAAGTTGTGAAAAAAGAGCTCGTCAATGATG    3437
CE1     GCTAATACAGTCTATATACCGCCATCTTCAGCAAACCATTTATGCAATGAAAGTTGTGAAAAAAGAGCTCGTCAATGATG    3430
CE2     GCTAATACAGTCTATATACCGCCATCTTCAGCAAACCATTTATGCTATGAAAGTTGTGAAAAAAGAGCTCGTCAATGATG    3437
GK1     GCTAATACAGTCTATATACCGCCATCTTCAGCAAACCATTTATGCAATGAAAGTTGTGAAAAAAGAGCTCGTCAATGATG    3429
GK2     GCTAATACAGTCTATATACCGCCATCTTCAGCAAACCATTTATGCAATGAAAGTTGTGAAAAAAGAGCTCGTCAATGATG    3437
LK1     GCTAATACAGTCTATATACCGCCATCTTCAGCAAACCATTTATGCAATGAAAGTTGTGAAAAAAGAGCTCGTCAATGATG    3430
CN1     GCTAATACAGTCTATATACCGCCATCTTCAGCAAACCATTTATGCAATGAAAGTTGTGAAAAAAGAGCTTGTCAATGATG    3429
CN2     GCTAATACAGTCTATATACCGCCATCTTCAGCAAACCATTTATGCAATGAAAGTTGTGAAAAAAGAGCTTGTCAATGATG    3437
MN1     GCTAATACAGTCTATATACCGCCATCTTCAGCAAACCATTTATGCAATGAAAGTTGTGAAAAAAGAGCTCGTCAATGATG    3437
MN2     GCTAATACAGTCTATATACCGCCATCTTCAGCAAACCATTTATGCAATGAAAGTTGTGAAAAAAGAGCTCGTCAATGATG    3437
B1      GCTAATACAGTCTATATACCGCCATCTTCAGCAAACCATTTATGCAATGAAAGTTGTGAAAAAAGAGCTTGTCAATGATG    3402
B2      GCTAATACAGTCTATATACCGCCATCTTCAGCAAACCATTTATGCAATGAAAGTTGTGAAAAAAGAGCTTGTCAATGATG    3437
SIT1    GCTAATACAGTCTATATACCGCCATCTTCAGCAAACCATTTATGCAATGAAAGTTGTGAAAAAAGAGCTTGTCAATGATG    3386
BB1     GCTAATACAGTCTATATACCGCCATCTTCAGCAAACCATTTATGCAATGAAAGTTGTGAAAAAAGAGCTCGTCAATGATG    3428
BB2     GCTAATACAGTCTATATACCGCCATCTTCAGCAAACCATTTATGCAATGAAAGTTGTGAAAAAAGAGCTCGTCAATGATG    3437
IMP     GCTAATACAGTCTATATACCGCCATCTTCAGCAAACCATTTATGCAATGAAAGTTGTGAAAAAAGAGCTTGTCAATGATG    3412
COW     GCTAATACAGTCTATATACCGCCATCTTCAGCAAACCATTTATGCAATGAAAGTTGTGAAAAAAGAGCTTGTCAATGATG    3419
BUF     GCTAATACAGTCTATATACCGCCATCTTCAGCAAACCATTTATGCAATGAAAGTTGTGAAAAAAGAACTTGTCAATGATG    3420
NIL     GCTAATACAGTCTATATACCGCCATCTTCAGCAAACCATTTATGCAATGAAAGTTGTGAAAAAAGAGCTTGTCAATGATG    3431

                3450       3460       3470       3480       3490       3500       3510       3520
                 .          .          .          .          .          .          .          .
GE1     ATGAGGTAAGCACAATGATGTTTTATTACCTCCAAACTATTACATTGGCTTAAGCATATGATATCAGGAAAAATGTCAGT    3517
CE1     ATGAGGTAAGCACAATGATGTTTTATTACCTCTAAACTATTACATTGGCTTAAGCATATGATATCAGGAAAAATGTCAGT    3510
CE2     ATGAGGTAAGCACAATGATGTTTTATTACCTCTAAACTATTACATTGGCTTAAGCATATGATATCAGGAAAAATGTCAGT    3517
GK1     ATGAGGTAAGCACAATGATGTTTTATTACCTCTAAACTATTACATTGGCTTAAGCATATGATATCAGGAAAAATGTCAGT    3509
GK2     ATGAGGTAAGCACAATGATGTTTTATTACCTCTAAACTATTACATTGGCTTAAGCATATGATATCAGGAAAAATGTCAGT    3517
LK1     ATGAGGTAAGCACAAT----------------------------------TTAAGCATATGATATCAGGAAAAATGTCAGT    3477
CN1     ATGAGGTAAGCACAATGATGTTTTATTACCTCTAAACTATTACATTGGCTTAAGCATATGATATCAGGAAAAATGTCAGT    3509
CN2     ATGAGGTAAGCACAATGATGTTTTATTACCTCTAAACTATTACATTGGCTTAAGCATATGATATCAGGAAAAATGTCAGT    3517
MN1     ATGAGGTAAGCACAATGATGTTTTATTACCTCTAAACTATTACATTGGCTTAACCATATGATATCAGGAAAAATGTCAGT    3517
MN2     ATGAGGTAAGCACAATGATGTTTTATTACCTCTAAACTATTACATTGGCTTAACCATATGATATCAGGAAAAATGTCAGT    3517
B1      ATGAGGTAAGCACAATGATGTTTTATTACCTCTAAACTATTACATTGGCTTAAGCGTATGATATCAGGAAAAATGTCAGT    3482
B2      ATGAGGTAAGCACAATGATGTTTTATTACCTCTAAACTATTACATTGGCTTAAGCGTATGATATCAGGAAAAATGTCAGT    3517
SIT1    ATGAGGTAAGCACAATGATGTTTTATTACCTCTAAACTATTACATTGGCTTAAGCATATGATATCAGGAAAAATGTCAGT    3466
BB1     ATGAGGTAAGCACAATGATGTTTTATTACCTCTAAACTATTACATTGGCTTAAGCGTATGATATCAGGAAAAATGTCAAT    3508
BB2     ATGAGGTAAGCACAATGATGTTTTATTACCTCTAAACTATTACATTGGCTTAAGCGTATGATATCAGGAAAAATGTCAAT    3517
IMP     ATGAGGTAAGCACAATGATGTTTTATTACCTCTAAATTATTACATTGGCTTAAACGTATGATATCAGAAAAAATGTCAAT    3492
COW     ATGAGGTAAGCACAATGATGTTTTATTACCTCTAAACTATTACATTGGCTTAAGCGTATGATATCAGAAAAGATGTCAAT    3499
BUF     ATGAGGTAAGCACAATGATGTTTTATTACCTCTAAACTATTACATTGGCTTAAGCGTATGATATCAGAAAAAATGTCAAT    3500
NIL     ATGAGGTAAGCACAATGATGTTTTATTACCTCTAAACTGTTACATTGGCTTAAGCATATGATATCAGAAAAAATGTCAAT    3511

                3530       3540       3550       3560       3570       3580       3590       3600
                 .          .          .          .          .          .          .          .
GE1     GTTTGTGACTTATTGTTTTAGTCTGCTTTTCTAAATACATGCTTCAGACTGCTTAATATACAAGATATATTTAAAGTTGT    3597
CE1     GTTTGTGACTTATTGTTTTAGTCTGCTTTTCTAAATACATGCTTCAGACTGCTTAATATACAAGATATATTTAAAGTTGT    3590
CE2     GTTTGTGACTTATTGTTTTAGTCTGCTTTTCTAAATACATGCTTCAGACTGCTTAATATACAAGATATATTTAAAGTTGT    3597
GK1     GTTTGTGACTTATTGTTTTAGTCTGCTTTTCTAAATACATGCTTCAGACTGCTTAATATACAAGATATATTTAAAGTTGT    3589
GK2     GTTTGTGACTTATTGTTTTAGTCTGCTTTTCTAAATACATGCTTCAGACTGCTTAATATACAAGATATATTTAAAGTTGT    3597
LK1     GTTTGTGACTTATTGTTTTAGTCTGCTTTTCTAAATACATGCTTCAGACTGCTTAATATACAAGATATATTTAAAGTTGT    3557
CN1     GTTTGTGACTTATTGTTGTAGTCTGCTTTTCTAAATACATGCTTCAGACTGCTTAATATACAAGATATATTTAAAGTTGT    3589
CN2     GTTTGTGACTTATTGTTGTAGTCTGCTTTTCTAAATACATGCTTCAGACTGCTTAATATACAAGATATATTTAAAGTTGT    3597
MN1     GTTTGTGACTTATTGTTTTAGTCTGCTTTTCTAAATACATGCTTCAGACTGCTTAATATACAAGATATATTTAAAGTTGT    3597
MN2     GTTTGTGACTTATTGTTTTAGTCTGCTTTTCTAAATACATGCTTCAGACTGCTTAATATACAAGATATATTTAAAGTTGT    3597
B1      GTTTGTGACTTATTGTTTTAGTCTGCTTTTCTAAATACATGCTTCAGACTGCTTAATATACAAGATATATTTAAAGTTGT    3562
B2      GTTTGTGACTTATTGTTTTAGTCTGCTTTTCTAAATACATGCTTCAGACTGCTTAATATACAAGATATATTTAAAGTTGT    3597
SIT1    GTTTGTGACTTATTGTTTTAGTCTGCTTTTCTAAATACATGCTTCAGACTGCTTAATATACAAGATATATTTAAAGTTGT    3546
BB1     GTTTGTGACTTATTGTTTTAGTCTGCTTTTCTAAATACATGCTTCAGACTGCTTAATATACAAGATATATTTAAAGTTGT    3588
BB2     GTTTGTGACTTATTGTTTTAGTCTGCTTTTCTAAATACATGCTTCAGACTGCTTAATATACAAGATATATTTAAAGTTGT    3597
IMP     GTTTGTGCCTTATCGTTTTTGTCTGCTTTTCTAAATACATGCTTCAGACTGCTTAATATATAAGATATATTTAAAGTTGT    3572
COW     GTTTGTGCCTTATTGTTTTAGTCTGCTTTTCTAAATACATGCTTCAGACTGCTTAAAATA-----TATATTTAAAGTTGT    3574
BUF     GTTTGTGCCTTATTGTTTTAGTCTGCTTTTCTAAATATATGCTTCAAACTGCTTAAAATA-----TATATTTAAAGTTGT    3575
NIL     GTTTGTGCCTTATTGTTTTAGTCTGCTTTTCTAAATACATGCTTCAGATTGCTTAATATATAAGATATATTTAAAGTTGT    3591
```

```
                3610       3620       3630       3640       3650       3660       3680
                 .          .          .          .          .          .          .
GE1     CTGATGATTATTAGATTTAAAAAAAAATAAACATCAAACCATATTATTACCAAGGGAAAGGACTTGAGCTTTGTAAGAGT    3677
CE1     CTGATGATTATTAGATTTAAAAAAAAATAAACATCAAACCATATTATTACCAAGGGAAAGGACTTGAGCTTTGTAAGAGT    3670
CE2     CTGATGATTATTAGATTTAGAAAAAAATAAACATCAAACCATATTATTACCTAGGGAAAGGACTTGAGCTTTGTAAGAGT    3677
GK1     CTGATGATTATTAGATTTAAAAAAAAATAAACATCAAACCATATTATTACCAAGGGAAAGGACTTGAGCTTTGTAAGAGT    3669
GK2     CTGATGATTATTAGATTTAAAAAAAAATAAACATCAAACCATATTATTACCAAGGGAAAGGACTTGAGCTTTGTAAGAGT    3677
LK1     CTGATGATTATTAGATTTAAAAAAAAATAAACATCAAACCATATTATTACCAAGGGAAAGTACTTGAGCTTTGTAAGAGT    3637
CN1     CTGATGATTATTAGATTTAAAAAAAAATAAACATCAAACCATATTATTACCAAGGGAAAGTACTTGAGCTTTGTAAGAGT    3669
CN2     CTGATGATTATTAGATTTAAAAAAAAATAAACATCAAACCATATTATTACCAAGGGAAAGTACTTGAGCTTTGTAAGAGT    3677
MN1     CTGATGATTATTAGATTTTTAAAAAAAATAAACATCAAACCATATTATTACCAAGGGAAAGGACTTGAGCTTTGTAAGAGT    3677
MN2     CTGATGATTATTAGATTTTTAAAAAAATAAACATCAAACCATATTATTACCAAGGGAAAGGACTTGAGCTTTGTAAGAGT    3677
B1      CTGATGATTATTAGATTTAAAAAAAAATAAACATCAAACCATATTATTACCAAGGGAAAGGACTTGAGCTTTGTAAGAGT    3642
B2      CTGATGATTATTAGATTTAAAAAAAAATAAACATCAAACCATATTATTACCAAGGGAAAGGACTTGAGCTTTGTAAGAGT    3677
SIT1    CTGATGATTATTAGATTTAAAAAAAAATAAACATCAAACCATATTATTACCAAGGGAAAGGACTTGAGCTTTGTAAGAGT    3626
BB1     CTGATGATTATTAGATTTAAAAAAAAATAAACATCAAACCATATTATTACCAAGGGAAAGTACTTGAGCTTTGTAAGAGT    3668
BB2     CTGATGATTATTAGATTTAAAAAAAAATAAACATCAAACCATATTATTACCAAGGGAAAGTACTTGAGCTTTGTAAGAGT    3677
IMP     CTAATGATTATTAGATT-AAAAAAAAATAAACATCAAACCATATTATTATCAGGGGAGAGTACTTGAGCTTTGTAAAAGT    3651
COW     CTGATGATTATTAGATT-AAAAAAAAATAAACATCAAACCATATTATTATCAAGGGAAAGTACTTGAGCTTTGTAAAAGT    3653
BUF     CTGATGATTATTAGATT-AAAAAAAAA--AACATCAAACCATATTATTATCAAGGGAAAGTACTTGAGCTTTGTAAAAGT    3652
NIL     CTGATGATTATTAGATTTAAAAAAAAATAAACATCAAACCATATTATTACCAAGGGAAAGTACTTGAGCTTTGTAAAAGT    3671

                3690       3700       3710       3720       3730       3740       3750       3760
                 .          .          .          .          .          .          .          .
GE1     TTGCTAGACTCTTTAATAATATTGAAAGATTTGCTGGGTAGTAGTTTCTCTTGTATATAAATTATTATTTCATGATTATC    3757
CE1     TTGCTAGACTCTTTAATAATATTGAAAGATTTGCTGGGTAGTAGTTTCTCTTGTATATAAATTATTATTTCATGATTATC    3750
CE2     TTGCTAGACTCTTTAATAATATTGAAAGATTTGCTGGGTAGTAGTTTCTCTTGTATATAAATTATTATTTCATGATTATC    3757
GK1     TTGCTAGACTCTTTAATAATATTGAAGGATTTGCTGGGTAGTAGTTTCTCTTGTATATAAATTATTATTTCATGATTATC    3749
GK2     TTGCTAGACTCTTTAATAATATTGAAGGATTTGCTGGGTAGTAGTTTCTCTTGTATATAAATTATTATTTCATGATTATC    3757
LK1     TTGCTAGACTCTTTAATAATATTGAAAGATTTGCTGGGTAATAGTTTCTCTTGTATATAAATTATTATTTCATGATTATC    3717
CN1     TTGCTAGACTCTTTAATAATATTGAAAGATTTGCTGGGTAGTAGTTTCTCTTGTATATAAATTATTATTTCATGATTATC    3749
CN2     TTGCTAGACTCTTTAATAATATTGAAAGATTTGCTGGGTAGTAGTTTCTCTTGTATATAAATTATTATTTCATGATTATC    3757
MN1     TTGCTAGACTCTTTAATAATATTGAAAGATTTGCTGGGTAGTAGTTTCTCATGTATATAAATTATTATTTCATGATTATC    3757
MN2     TTGCTAGACTCTTTAATAATATTGAAAGATTTGCTGGGTAGTAGTTTCTCATGTATATAAATTATTATTTCATGATTATC    3757
B1      TTGCTAGACTCTTTAATAATATTGAAAGATTTGCTGGCTAGTAGTTTCTCTTGTATATAAATTATTATTTCATGATTATC    3722
B2      TTGCTAGACTCTTTAATAATATTGAAAGATTTGCTGGGTAGTAGTTTCTCTTGTATATAAATTATTATTTCATGATTATC    3757
SIT1    TTGCTAGACTCTTTAATAATATTGAAAGATTTGCTGGGTAGTAGTTTCTCTTGTATATAAATTATTATTTCATGATTATC    3706
BB1     TTGCTAGACTCTTTAATAATATTGAAAGATTTGCTGGGTAGTAGTTTCTCTTGTATATAAATTATTATTTCATGATTATC    3748
BB2     TTGCTAGACTCTTTAATAATATTGAAAGATTTGCTGGGTAGTAGTTTCTCTTGTATATAAATTATTATTTCATGATTATC    3757
IMP     TTGCTAGACTCTTTAATAATATTGAAAGATTTGCTGGGTAGTAGTTTCTCTTGTATATAAATTATTATTTCATGATCATC    3731
COW     TTGGTAGACTCTTTAATAATATTGAAAGGTTTGCTGGGTAGTAGTTTCTCTTGTATATAAATTATTATTTCATGATTATC    3733
BUF     TTGCTAGACTCTTTAATAATATTGAAAGATTTGCTGGGTAGTAGTTTCTCTTGTATATAAATTATTATTTCATGATTATC    3732
NIL     TTGCTAGACTCTTTAATAATATTGAAAGATTTGCTGGGTAGTAGTTTCTCTTGTATATAAATTATTATTTCATGATTATC    3751

                3770       3780       3790       3800       3810       3820       3830       3840
                 .          .          .          .          .          .          .          .
GE1     AATGCTTTTGATGTTTAGCCTAGTTAGTCAATAATTGCATTTATGGTAATAGAATTACTAAGAACAAATTTGTCATAGTT    3837
CE1     AATGCTTTTGATGTTTAGCCTAGTTAGTCAATAATTGCATTTATGGTAATAGAATTACTAAGAACAAATTTGTCATAGTT    3830
CE2     AATGCTTTTGATGTTTAGCCTAGTTAGTCAATAATTGCATTTATGGTAATAGAATTACTAAGAACAAACTTGTCATAGTT    3837
GK1     AATGCTTTTGACGTTTAGCCTAGTTAGTCAATAATTGCATTTATGGTAATAGAATTACTAAGAACAAATTTGTCATAGTT    3829
GK2     AATGCTTTTGACGTTTAGCCTAGTTAGTCAATAATTGCATTTATGGTAATAGAATTACTAAGAACAAATTTGTCATAGTT    3837
LK1     AATGCTTTTGATGTTTAGCCTAGTTAGTCAATAATTGCATTTATGGTAATAGAATTACTAAGAACAAATTTGTCATAGTT    3797
CN1     AATGCTTTTGATGTTTAGCCTAGTTAGTCAATAATTGCATTTATGGTAATAGAATTACTAAGAACAAATTTGTCATAGTT    3829
CN2     AATGCTTTTGATGTTTAGCCTAGTTAGTCAATAATTGCATTTATGGTAATAGAATTACTAAGAACAAATTTGTCATAGTT    3837
MN1     AATGCTTTTGATGTTTAGCCTAGTTAGTCAATAATTGCATTTATGGTAATAGAATTACTAAGAACAAATTTGTCATAGTT    3837
MN2     AATGCTTTTGATGTTTAGCCTAGTTAGTCAATAATTGCATTTATGGTAATAGAATTACTAAGAACAAATTTGTCATAGTT    3837
B1      AATGCTTTTGATGTTTAGCCTAGTTAGTCAATAATTGCATTTATGGTAATAGAATTACTAAGAACAAATTTGTCATAGTT    3802
B2      AATGCTTTTGATGTTTAGCCTAGTTAGTCAATAATTGCATTTATGGTAATAGAATTACTAAGAACAAATTTGTCATAGTT    3837
SIT1    AATGCTTTTGATGTTTAGCCTAGTTAGTCAATAATTGCATTTATGGTAATAGAATTACTAAGAACAAATTTGTCATAGTT    3786
BB1     AATGCTTTTGATGTTTAGCCTAGTTAGTCAATAATTGCATTTATGGTAATAGAATTACTAAGAACAAATTTGTCATAGTT    3828
BB2     AATGCTTTTGATGTTTAGCCTAGTTAGTCAATAATTGCATTTATGGTAATAGAATTACTAAGAACAAATTTGTCATAGTT    3837
IMP     AATGCTTGTGATGTTTAGCCCAGTTAGTCAATAGTTGCATTTATGGTAATAGAATTACTAAGAACAAATTTGTCATAGTC    3811
COW     AATGCTTTTGATGTTTAGCCTAGTTAGT----AATTGCATTTATGGTAATAGAATTACTAAGAACAAATTTGTCATAGTT    3809
BUF     AATGCTTTTGATGTTTAGCCTAGTTAGT----AATTGCATTTATGGTAATAGAATTACTAAGAACAAATTTGTCATAGTT    3808
NIL     AGTGCTTTTGATGTTTAGCCTAGTTAGTCAATAATTGCATTTATGGTAATAGAATTACTAAGAACAAATTTGTCATAGTT    3831
```

```
                3850       3860       3870       3880       3890       3900       3910       3920
                .          .          .          .          .          .          .          .
GE1     TTGATACTGTCATTCTAGTTTGACATAAATTTCTTTCATATTTTAAAATAGGATCATTCCTAATATAAATGTCAAAAGCA    3917
CE1     TTGATACTGTCATTCTAGTTTGACATAAATTTCTTTCATATTTTAAAATAGGATCATTCCTAATATAAATGTCAAAAGCA    3910
CE2     TTGATACTGTCATTCTAGTTTGACATAAATTTCTTTCATATTTTAAAATAGGATCATTCCTAATATAAATGTCAAAAGCA    3917
GK1     TTGATACTGTCATTCTAGTTTGACATAAATTTCTTTCATATTTTAAAATAGGATCATTCCTAATATAAATGTCAAAAGCA    3909
GK2     TTGATACTGTCATTCTAGTTTGACATAAATTTCTTTCATATTTTAAAATAGGATCATTCCTAATATAAATGTCAAAAGCA    3917
LK1     TTGATACTGTCATTCTAGTTTGTCATAAATTTCTTTCATATTTTAAAATAGGATCATTCCTAATATAAATGTCAAAAGCA    3877
CN1     TTGATACTGTCATTCTAGTTTGACATAAATTTCTTTCATATTTTAAAATAGGATCATTCCTAATATAAATGTCAAAAGCA    3909
CN2     TTGATACTGTCATTCTAGTTTGACATAAATTTCTTTCATATTTTAAAATAGGATCATTCCTAATATAAATGTCAAAAGCA    3917
MN1     TTGATACTGTCATTCTAGTTTGACATAAATTTCTTTCATATTTTAAAATAGGATCATTCCTAATATAAATGTCAAAAGCA    3917
MN2     TTGATACTGTCATTCTAGTTTGACATAAATTTCTTTCATATTTTAAAATAGGATCATTCCTAATATAAATGTCAAAAGCA    3917
B1      TTGATACTGTCATTCTAGTTTGACATAAATTTCTTTCATATTTTAAAATAGGATCATTCCTAATATAAATGTCAAAAGCA    3882
B2      TTGATACTGTCATTCTAGTTTGACATAAATTTCTTTCATATTTTAAAATAGGATCATTCCTAATATAAATGTCAAAAGCA    3917
SIT1    TTGATACTGTCATTCTAGTTTGACATAAATTTCTTTCATATTTTAAAATAGGATCATTCCTAATATAAATGTCAAAAGCT    3866
BB1     TTGATACTGTCATTCTAGTTTGACATAAATTTCTTTCATATTTTAAAATAGGATCATTCCTAATATAAATGTCAAAAGCA    3908
BB2     TTGATACTGTCATTCTAGTTTGACATAAATTTCTTTCATATTTTAAAATAGGATCATTCCTAATATAAATGTCAAAAGCA    3917
IMP     TTGATACTGTCATTCTAGTTTGAGATAAATTTCTTTCATATTTTAAAATAGGATCATTCCTAATATAAATGTCAAAAGCA    3891
COW     TTGATACTGTCATTCTAGTTTGACATAAATTTCTTTCATATTTTAAAATAGGATCATTCCTAATATAAATGTAGAAAGCA    3889
BUF     TTGATACTGTTATTCTAGTTTGACATAAATTTCTTTCATATTTTAAAATAGGATCATTCCTAATATAAATGTAAAAAGCA    3888
NIL     TTGATACTGTCATTCTAGTTTGACATAAATTTCTTTCATATTTTAAAATAGGATCATTCCTAATATAAATATAGAAAGCA    3911

                3930       3940       3950       3960       3970       3980       3990       4000
                .          .          .          .          .          .          .          .
GE1     AGACTGAATCTTATATGTGTAGGTAAGAATCTACCTCTGACTCAAATATATATGTAGCTTGAAACAGACATGTGTGTGT-    3996
CE1     AGACTGAATCTTATATGTGTAGGTAAGAATCTACCTCTGACTCAAATATATATGTAGCTTGAAACAGACATGTGTGTGT-    3989
CE2     AGACTGAATCTTATATGTGTAGGTAAGAATCTACCTCTGACTCAAATATATATGTAGCTTGAAACAGACATGTGTGTGT-    3996
GK1     AGACTGAATCTTAGATGTGTAGGTAAGAATCTACCTCTGACTCAAATATATATGTAGCTTGAAACAGACACGTGTGTGT-    3988
GK2     AGACTGAATCTTAGATGTGTAGGTAAGAATCTACCTCTGACTCAAATATATATGTAGCTTGAAACAGACACGTGTGTGT-    3996
LK1     AGACTGAATCTTATATGTGTAGGTAAGAATCTACCTCTGACTCAAATATATATGTAGCTTGAAACAGACACGTGTGTGT-    3956
CN1     AGACTGAATCTTATATGTGTAGGTAAGAATCTACCTCTGACTCAAATATATATGTAGCTTGAAACAGACACGTGTGTGT-    3988
CN2     AGACTGAATCTTATATGTGTAGGTAAGAATCTACCTCTGACTCAAATATATATGTAGCTTGAAACAGACACGTGTGTGT-    3996
MN1     AGACTGAATCTTATATGTGTAGGTAAGAATCTACCTCTGACTCAAATATATATGTAGCTTGAAACAGACACGTGTGTGT-    3996
MN2     AGACTGAATCTTATATGTGTAGGTAAGAATCTACCTCTGACTCAAATATATATGTAGCTTGAAACAGACACGTGTGTGT-    3996
B1      AGACTGAATCTTATATGTGTAGGTAAGAATCTACCTCTGACTCAAATATATATGTAGCTTGAAACAGACACGTGTGTGT-    3961
B2      AGACTGAATCTTATATGTGTAGGTAAGAATCTACCTCTGACTCAAATATATATGTAGCTTGAAACAGACACGTGTGTGT-    3996
SIT1    C-ACTGAATCTTATATGTGTAGGTAAGAATCTACCTCTGACTCAAATATATATGTAGCTTGAAACGGACACGTGTGTGT-    3944
BB1     AGACTGAATCTTATATGTGTAGGTAAGAATCTACCTCTGACTCAAATATATATGTAGCTTGAAACAGACACGTGTGTGT-    3987
BB2     AGACTCAATCTTATATGTGTAGGTAAGAATCTACCTCTGACTCAAATATATATGTAGCTTGAAACAGACACGTGTGTGT-    3996
IMP     AGACTGAATC--ATATGTGTAGGTAAGAATCTACCTCTGACTCAAATATATATGTAGCTAGAAAAAGAC----GTGCGT-    3964
COW     AGACTGAATCTTTTATGTATAGGTAAGAATCTACCTCTGACTCAAATATATATGTAGCTAGAAACAGACATGTGTGTGTA    3969
BUF     AGACTGAATCTTATATGTGTAGGTAAGAATCTACCTCTGACTCAAATATATATGTAGCTAGAAACAGAC--GTGTGTGT-    3965
NIL     AGACTGAATCTTATATGTGTAGGTAAGAATCTACCTCTGACTCAAATATATATGTAGCTAGAAAAAGAC--GTGTGTGT-    3988

                4010       4020       4030       4040       4050       4060       4070       4080
                .          .          .          .          .          .          .          .
GE1     -GTATGTCTAATATAAAAGCAAATTATCGGTGGCCTGGAAAAAAAA--------GTTATATATTTGTGGGTGTTCAAGCC    4067
CE1     -GTATGTCTAATATAAAAGCAAATTATCGGTGGCATGAAAAAAAA---------GTTATGTATTTGTTGGTGTTCAAGCC    4059
CE2     -GTATGTCTAATATAAAAGCAAATTATCGGTGGCATGAAAAAAAA---------GTTATATATTTGTTGGTGTTCAAGCC    4066
GK1     -GTATGTCTAATATAAAAGCAAATTATCAGTGGCATGAAAAAAAAAA-------GTTATATATTTGTTGGGGTTCAAGCC    4060
GK2     -GTATGTCTAATATAAAAGCAAATTATCAGTGGCATGAAAAAAAAAA-------GTTATATATTTGTTGGTGTTCAAGCC    4068
LK1     -GTATGTCTAATATAAAAGCAAATTATCAGTGGCATGAAAAAAA----------GTTATATATTTGTT------CAAGCC    4019
CN1     -GTATGTCTAATATAAAAGCAAATTATCAGTGGCATGAAAAAAAA---------GTTATATATTTGTTGGTGTTCAAGCC    4058
CN2     -GTATGTCTAATATAAAAGCAAATTATCAGTGGCATGAAAAAAAA---------GTTATATATTTGTTGGTGTTCAAGCC    4066
MN1     -GTATGTCTAATATAAAAGCAAATTATCAGTGGCATGAAAAAAAAA--------GTTATATATTTGTTGGTGTTCAAGCC    4067
MN2     -GTATGTCTAATATAAAAGCAAATTATCAGTGGCATGAAAAAAAAA--------GTTATATATTTGTTGGTGTTCAAGCC    4067
B1      -GTATGTCTAATATAAAAGCAAATTATCAGTGGCATGAAAAAAAAA--------GTTATATATTTGTTGGTGTTCAAGCC    4032
B2      -GTATGTCTAATATAAAAGCAAATTATCAGTGGCATGAAAAAAAAA--------GTTATATATTTGTTGGTGTTCAAGCC    4067
SIT1    -GTATGTCTACTATAAAAGCAAATTATCAGTGGCATGAAAAAAAAA-------GTTATATATTTGTTGGGGTTCAAGCC    4016
BB1     -GTATGTCTAATATAAAAGCAAATTATCAGTGGCATGAAAAAAAAA-------GTTATATATTTGTTGGTGTTCAAGCC    4059
BB2     -GTATGTCTAATATAAAAGCAAATTATCAGTGGCATGAAAAAAAAA-------GTTATATATTTGTTGGTGTTCAAGCC    4068
IMP     -ATATGTCTAATATAAAAGCAAAT-ATCAGTGGCATGAAAAAAAAA-----CAGTTACATATTTGTTGGTGTTCAAGCC    4037
COW     TGTATGTCTAATATAAAAGCAAAT-ATCAGTGGCATGAAAAAAAAA------CAGTTATATATTTGTTGGTGTTCAAGCC    4042
BUF     -GTATGTCTAATATAAAAGCAAAT-ATCAGTGGCATGAAAAAAAAAAAAAAACAGTTATATATTTGTTGGTGTTCAAGCC    4043
NIL     -GTATGTCTAATATAAAAGCAAAT-ATCAGTGGCATGAAAAAAAAA------CAGTTATATATTTGTTGGTGTTCAAGCC    4060
```

```
                4090       4100       4110       4120       4130       4140       4150       4160
                   .          .          .          .          .          .          .          .
GE1     CAAACAACACTTAACCATCTTCAAAGTACAGAATTACGCACGAGTAAACGATATATGAGATAGTCACTCATTTTCACCTG    4147
CE1     CAAACAACACTTAACCATCTTCAAAGTACAGAATTACGCACGAGTAAACGATATATGAGATAGTCACTCATTTTCACCTG    4139
CE2     CAAACAACACTTAACCATCTTCAAAGTACAGAATTACGCACGAGTAAACGATATATGAGATAGTCACTCATTTTCACCTG    4146
GK1     CAAACAACACTTAACCATCTTCAAAGTACAGAATTACGCACGAGTAAACGATATATGAGATAGTCACTCATTTTCACCTG    4140
GK2     CAAACAACACTTAACCATCTTCAAAGTACAGAATTACGCACGAGTAAACGATATATGAGATAGTCACTCATTTTCACCTG    4148
LK1     CAAACAACACTTAACCATCTTCAAAGTACAGAATTACTCACGAGTAAACGATATATGAGATAGTAACTCATTTTCACCTG    4099
CN1     CAAACAACACTTAACCATCTTCAAAGTACAGAATTACTCACGAGTAAACGATATATGAGATAGTAACTCATTTTCACCTG    4138
CN2     CAAACAACACTTAACCATCTTCAAAGTACAGAATTACTCACGAGTAAACGATATATGAGATAGTAACTCATTTTCACCTG    4146
MN1     CAAACAACACTTAACCATCTTCAAAGTACAGAATTACGCACGAGTAAACGATATATGAGATAGTCACTTATTTTCACCTG    4147
MN2     CAAACAACACTTAACCATCTTCAAAGTACAGAATTACGCACGAGTAAATGATATATGAGATAGTCACTTATTTTCACCTG    4147
B1      CAAACAACACTTAACCATCTTCAAAGTACAGAATTACGCACGAGTAAACGATATATGAGATAGTCACTCATTTTCACCTG    4112
B2      CAAACAACACTTAACCATCTTCAAAGTACAGAATTACGCACGAGTAAACGATATATGAGATAGTCACTCATTTTCACCTG    4147
SIT1    CAAACAACACTTAACCATCTTCAAAGTACAGAATTACGCACAAGTAAACGATATATGAGATAGTCACTCATTTTCACCTG    4096
BB1     CAAACAACACTTAACCATCTTCAAAGTACAGAATTACGCACGAGTAAATGATATATGAGATAGTCACCCATTTTCCCCTG    4139
BB2     CAAACAACACTTAACCATCTTCAAAGTACAGAATTACGCACGAGTAAATGATATATGAGATAGTCACTCATTTTCACCTG    4148
IMP     CAAACAACACTTAACCATCTTCAAAGTACAGAACTGTGCATGAGCAAATGATATATGAGAAAGTCACTCATTTTCACCTG    4117
COW     CAAACAATACTTAACCATCTTCAAAGTACAGAATTGCACATGAGTAAACGATATATGAGATAGTCACTCATTTTCACCTG    4122
BUF     CAAACAACACTTAACCATCTTCAAAGTACAGAATTGCACATGAGTAAACGATATATGAGATAGTCACTCATTTTCACCTG    4123
NIL     CAAACAACACTTAACCATCTTCAAAGTACAGAATTGCGCATGAGTAAATGATATATGAGATAGTCACTCATTTTCACCTG    4140

                4170       4180       4190       4200       4210       4220       4230       4240
                   .          .          .          .          .          .          .          .
GE1     ATATTTATTAGTGATATACAGTTATATTCTTCTTCACTTCCAGTCTTCTTAATGTCCCCTTCTCAAATCTTTTCCTTAGG    4227
CE1     ATATTTATTAGTAATATACAGTTATATTCTTCTTCACTTCCAGTCTTCTTAATGTCCCCTTCTCAAATCTTTTCCTTAGG    4219
CE2     ATATTTATTAGTAATATACAGTTATATTCTTCTTCACTTCCAGTCTTCTTAATGTCCCCTTCTCAAATCTTTTCCTTAGG    4226
GK1     ATATTTATTAGTAATGTACAGTTATATTCTTCTTCACTTCCAGTCTTCTTAATGTCCCCTTCTCAAATCTTTTCCTTAGG    4220
GK2     ATATTTATTAGTAATGTACAGTTATATTCTTCTTCACTTCCAGTCTTCTTAATGTCCCCTTCTCAAATCTTTTCCTTAGG    4228
LK1     ATATTTATTAGTAATATACAGTTATATTCTTCTTCACTTCCAGTCTTCTTAATGTCCCCTTCTCAAATCTTTTCCTTAGG    4179
CN1     ATATTTATTAGTAATATACAGTTATATTCTTCTTCACTTCCAGTCTTCTTAATGTCCCCTTCTCAAATCTTTTCCTTAGG    4218
CN2     ATATTTATTAGTAATATACAGTTATATTCTTCTTCACTTCCAGTCTTCTTAATGTCCCCTTCTCAAATCTTTTCCTTAGG    4226
MN1     ATATTTATTAGTAATGTACAGTTATATTCTTCTTCACTTCCAGTCTTCTTAATGTCCCCTTCTCAAATCTTTTCCTTAAG    4227
MN2     ATATTTATTAGTAATGTACAGTTATATTCTTCTTCACTTCCAGTCTTCTTAATGTCCCCTTCTCAAATCTTTTCCTTAAG    4227
B1      ATATTTATTAGTAATGTACAGTTATATTCTTCTTCACTTCCAGTCTTCTTAATGTCCCCTTCTCAAATCTTTTCCTTAGG    4192
B2      ATATTTATTAGTAATGTACAGTTATATTCTTCTTCACTTCCAGTCTTCTTAATGTCCCCTTCTCAAATCTTTTCCTTAGG    4227
SIT1    ATATTTATTAGTAATGTACAGTTATATTCTTCTTCACTTCCAGTCTTCTTAATGTCCCCTTCTCAAATCTTTTCCTTAGG    4176
BB1     ATATTTATTAGTAATGTACAGTTATATTCTTCTTCACTTCCAGTCTTCTTAATGTCCCCTTCTCAAATCTTTTCCTTAGG    4219
BB2     ATATTTATTAGTAATGTACAGTTATATTCTTCTTCACTTCCAGTCTTCTTAATGTCCCCTTCTCAAATCTTTTCCTTAGG    4228
IMP     ATATTCATTAGTAATATACAATTATATTCCTCTTCACTTCCAATCTTCTTAATGTGCCCTTCTCAAGTCTTTTCCCTAGG    4197
COW     ATATTTACTAGTAATATACAATTATATTCTTCTTCACTTCCAGTCTTCTTAATGTCCCCTTCTCAAATCTTTTCCTTAGG    4202
BUF     ATATTTACTAGTAATATACAATTATATTCTTCTTCACTTCCAGTCTTTTTAATGTCCCCTTCTCAAATCTTTTCCTTAGG    4203
NIL     ATATTTACTAGTAATATACAATTGTATTCTTCTTCACTTCCAGTCTTCTTAATGTCCCCTTTTCAAATCTTTTCCTTAGG    4220

                4250       4260       4270       4280       4290       4300       4310       4320
                   .          .          .          .          .          .          .          .
GE1     GGTCTAATTCTGAGTCATGAATTCTCATCTTCACCACTTACTCTGTACTGAGACTTCCTCCCGGCTACCATCCCTATGAT    4307
CE1     GGTCTAATTCTGAGTCATGAATTCTCATCTTCACCACTTACTCTGTACTGAGACTTCCTCCCGGCTACCATCCCTATGAT    4299
CE2     GGTCTAATTCTGAGTCATGAATTCTCATCTTCACCACTTACTCTGTACTGAGACTTCCTCCCGGCTACCATCCCTATGAT    4306
GK1     GGTCTAATTCTGAGTCATGAATTCTCATCTTCACCACTTACTCTGTACTGAAACTTCCTCTCAGCTACCATCCCTATGAT    4300
GK2     GGTCTAATTCTGAGTCATGAATTCTCATCTTCACCACTTACTCTGTACTGAAACTTCCTCTCAGCTACCATCCCTATGAT    4308
LK1     GGTCTAATTCTGAGTCATGAATTCTCATCTTCACCACTTACTCTGTACTGAGACTTCCTCCCGACTACCATCCCTGTGAT    4259
CN1     GGTCTAATTCTGAGTCATGAATTCTCATCTTCACCACTTACTCTGTACTGAGACTTCCTCCCGGCTACCATCCCTATGAT    4298
CN2     GGTCTAATTCTGAGTCATGAATTCTCATCTTCACCACTTACTCTGTACTGAGACTTCCTCCCGGCTACCATCCCTATGAT    4306
MN1     GGTCTAATTCTGAGTCATGAATTCTCATCTTCACCACTTACTCTGTACTGAGACTTCCTCTCAGCTACCATCCCTATGAT    4307
MN2     GGTCTAATTCTGAGTCATGAATTCTCATCTTCACCACTTACTCTGTACTGAGACTTCCTCTCAGCTACCATCCCTATGAT    4307
B1      GGTCTAATTCTGAGTCATGAATTCTCATCTTCACCACTTACTCTGTACTGAGACTTCCTCTCAGCTACCATCCCTATGAT    4272
B2      GGTCTAATTCTGAGTCATGAATTCTCATCTTCACCACTTACTCTGTACTGAGACTTCCTCTCAGCTACCATCCCTATGAT    4307
SIT1    GGTCTAATTCTGAGTCATGAACTCTCATCTTCACCACTTACTCTGTACTGAGACTTCCTCTCAGCTACCATCCCTATGAT    4256
BB1     GGTCTAATTCTGAGTCATGAATTCTCATCTTCACCACTTACTCTGTACTGAGACTTCCTCTCAGCTACCATCCCTATGAT    4299
BB2     GGTCTAATTCTGAGTCATGAATTCTCATCTTCACCACTTACTCTGTACTGAGACTTCCTCTCAGCTACCATCCCTATGAT    4308
IMP     GGTCTAACTCTGAGTTATGAATTCTCATCTTCACCACTTACCCTGTATTGAAACTTCCTCCCCACTACCATCTCTATGAT    4277
COW     GGTCTAATTCTGAGTCATGAATTCTCATCTTCATCACTTACCCTGTACTGAGACTTCCTCCCCGCTACCATCCCTATGAT    4282
BUF     GGTCTAATTCTGAGTCATGAATTCTCATCTTCACCACTTACCCTGTCCTGAGACTTCCTCCCTGCTACCATCCCTATGAT    4283
NIL     GGTCAAATTCTGAGTCATGAATTCTCATCTTCACCACTTACTCTGTACTGAGACTTCCTCCCTGCTACCATCCCTATGAT    4300
```

```
                4330       4340       4350       4360       4370       4380       4390       4400
                   .          .          .          .          .          .          .          .
GE1     ACGAAATTTTAATTTCATTACTTATTTTGAAAAAGTGATCAATTGCCAACCATTT-CATTATTTACATTTATTAGCATGT    4386
CE1     ACGAAATTTTAATTTCATTACTTATTTTGAAAAAGTGATCAATTGCCAACCATTT-CATTATTTACATTTATTAGCATGT    4378
CE2     ACGAAATTTTAATTTCATTACTTATTTTGAAAAAGTGATCAATTGCCAACCATTT-CATTATTTACATTTATTAGCATGT    4385
GK1     ACAAAATTTTAATTTCATTACTTATTTTGAAAAAGTGATCAATTGCCA-CCATTT-CATTATTTACATTTATTAGCATGT    4378
GK2     ACAAAATTTTAATTTCATTACTTATTTTGAAAAAGTGATCAATTGCCA-CCATTT-CATTATTTACATTTATTANCATGT    4386
LK1     ACGAAATTTTAATTTCATTACTTATTTTGAAAAAGTGATCAATTGCCAACCATTT-CATTATTTACATTTATTAGCATGT    4338
CN1     GCGAAATTTTAATTTCATTACTTATTTTGAAAAAGTGATCAATTGCCAACCATTT-CATTATTTACATTTATTAGCATGT    4377
CN2     GCGAAATTTTAATTTCATTACTTATTTTGAAAAAGTGATCAATTGCCAACCATTT-CATTATTTACATTTATTAGCATGT    4385
MN1     ACAAAATTTTAATTTCATTACTTATTTTGAAAAAGTGATCAATTGCCA-CCATTT-CATTATTTACATTTATTAGCATGT    4385
MN2     ACAAAATTTTAATTTCATTACTTATTTTGAAAAAGTGATCAATTGCCA-CCATTT-CATTATTTACATTTATTAGCATGT    4385
B1      ACAAAATTTTAATTTCATTACTTATTTTGAAAAAGTGATCAATTGCCA-CCATTT-CATTATTTACATTTATTAGCATGT    4350
B2      ACAAAATTTTAATTTCATTACTTATTTTGAAAAAGTGATCAATTGCCA-CCATTT-CATTATTTACATTTATTAGCATGT    4385
SIT1    ACAAAATTTTAATTTCATTACTTATTTTGAAAAAGTGATGAATTGCCA-CCATTT-CATTATTTACATTTATTAGCATGT    4334
BB1     ACAAAATTTTAATTTCATTACTTATTTTGAAGAAGTGATCAATTGCCA-CCATTT-CATTATTTACATTTATTAGCATGT    4377
BB2     ACAAAATTTTAATTTCATTACTTATTTTGAAAAAGTGATCAATTGCCA-CCATTT-CATTATTTACATTTATTAGCATGT    4386
IMP     ACAACATTTTAATTTCATTACTTATTTTGAAAAAGTGACCAATTGCCAACCATTTGCATTATTTACATTTATTAGCATGT    4357
COW     ATGAAATTTTAATTTCATTACTTATTTTGAAAAAGTGATCAATTGCCAACCATTTGCATTATTTACATTTGTT-------    4355
BUF     ACGAAATTTTAATTTCAGTACTTATTTTGAAAAAGTGACCAATTGCCAACCATTTGCATTATTTACATTTATGAGCGTGT    4363
NIL     ACGAAATTTTAATTTCATCACTTATTTTGAAAAAGTGACGAATTGCCAACCATTTGCATTATTTACATTTATTAGCATGT    4380

                4410       4420       4430       4440       4450       4460       4470       4480
                   .          .          .          .          .          .          .          .
GE1     GTAGGCAGTGAATCTTTGTTCTGTAAGGTGACTTTGGTGAATTTCTTAATAAAAATAGAATAGCAGGACCCAGTATAAAA    4466
CE1     GTAGGCAGTGAATCTTTGTTCTGTAAGGTGACTTTGGTGAATTTCTTAATAAAAATAGAATAGCAGGACCCAGTATAAAA    4458
CE2     GTAGGCAGTGAATCTTTGTTCTGTAAGGTGACTTTGGTGAATTTCTTAATAAAAATAGAATAGCAGGACCCAGTATAAAA    4465
GK1     GTAGGCAGTGAATCTTTGTTCTGTAAGGGGACTTTGGTGAATTTCTTAATAAAAATAGAATAGCAGGACCCAGTATAAAA    4458
GK2     GTAGGCAGTGAATCTTTGTTCTGTAAGGGGACTTTGGTGAATTTCTTAATAAAAATAGAATAGCAGGACCCAGTATAAAA    4466
LK1     GTAGGCAGTGAATCTTTGTTCTGTAAGGTGACTTTGGTGAATTTCTTAATAAAAATAGAATAGCAGGACCCAGTATAAAA    4418
CN1     GTAGGCAGTGAATCTTTGTTCTGTAAGGTGACTTTGGTGAATTTCTTAATCAAAATAGAATAGCAGGACCCAGTATAAAA    4457
CN2     GTAGGCAGTGAATCTTTGTTCTGTAAGGTGACTTTGGTGAATTTCTTAATCAAAATAGAATAGCAGGACCCAGTATAAAA    4465
MN1     GTAGGCAGTGAATCTTTGTTCTGTAAGGTGACTTTGGTGAATTTCTTAATAAAAATAGAATAGCAGGACCCAGTATAAAA    4465
MN2     GTAGGCAGTGAATCTTTGTTCTGTAAGGTGACTTTGGTGAATTTCTTAATAAAAATAGAATAGCAGGACCCAGTATAAAA    4465
B1      GTAGGCAGTGAATCTTTGTTCTGTAAGATGACTTTGGTGAATTTCTTAATAAAAATAGAATAGCAGGACCCAGTATAAAA    4430
B2      GTAGGCAGTGAATCTTTGTTCTGTAAGATGACTTTGGTGAATTTCTTAATAAAAATAGAATAGCAGGACCCAGTATAAAA    4465
SIT1    GTAGGCAGTGAATCTTTGTTCTGTAAGGTGACTTTGGTGAATTTCTTAATAAAAATAGAATAGCAGGACCCAGTATAAAA    4414
BB1     GTAGGCGGTGAATCTTTGTTCTGTAAGGTGACTTTGGTGAATTTCTTAATAAAAATAGAATAGCAGGACCCAGTATAAAA    4457
BB2     GTAGGCGGTGAATCTTTGTTCTGTAAGGTGACTTTGGTGAATTTCTTAATAAAAATAGAATAGCAGGACCCAGTATAAAA    4466
IMP     GTAGGCAGTGAATCTTTGTTCTATAAGGTGACTTTGGTGAATTTCTTAATAAAAATAGAATAGTAGGACCCAGTATAAAA    4437
COW     ----------------------TAAAGTGACTTTGGTGAATTTCTTAATAAAAATAGAAAAGCAGGACCAAGTATAAAA    4412
BUF     GTAGGCAGTGAATCTTTGTTCTGTAAGGTGGCTTTGGTGAATTTCTTAATAAAAATAGAAAAGCAGGACCCAGTATAAAA    4443
NIL     GTAGGCAGTGAATCTTTGTTGTGTAAGGTGACTTTGGTGAATTTCTTAATAAAAATAGAATAGCAGGACCCAGTATAAAA    4460

                4490       4500       4510       4520       4530       4540       4550       4560
                   .          .          .          .          .          .          .          .
GE1     TAAACAGGATCTGAATGATCCAGATTTGGGAAGCAGGCCTTTTAAGTACTCTAATATTGCACTTCACCTTTTGAGGACTC    4546
CE1     TAAACAGGATCTGAATGATCCAGATTTGGGAAGCAGGCCTTTTAAGTACTCTAATATTGCACTTCACCTTTTGAGGACTC    4538
CE2     TAAACAGGATCTGAATGATCCAGATTTGGGAAGCAGGCCTTTTAAGTACTCTAATATTGCACTTCACCTTTTGAGGACTC    4545
GK1     TAAACAGGATCTGAATGATCCAGATTTGGGAAGCAGGCCTTTTAAGAACTCTAATATTGCACTTCACCTTTTGAGGACTC    4538
GK2     TAAACAGGATCTGAATGATCCAGATTTGGGAAGCAGGCCTTTTAAGTACTCTAATATTGCACTTCACCTTTTGAGGACTC    4546
LK1     TAAACAGGATCTGAATGATCCAGATTTGGGAAGCAGGCCTTTTAAGTACTCTAATATTGCACTTCACCTTTTGAGGACTC    4498
CN1     TAAACAGGATCTGAATGATCCAGATTTGGGAAGCAGGCCTTTTAAGTACTCTAATATTGCACTTCACCTTTTGAGGACTC    4537
CN2     TAAACAGGATCTGAATGATCCAGATTTGGGAAGCAGGCCTTTTAAGTACTCTAATATTGCACTTCACCTTTTGAGGACTC    4545
MN1     TAAACAGGATCTGAATGATCCAGATTTGGGAAGCAGGCCTTTTAAGTACTCTAATATTGCACTTC-CCTTTTGAGGACTC    4544
MN2     TAAACAGGATCTGAATGATCCAGATTTGGGAAGCAGGCCTTTTAAGTACTCTAATATTGCACTTC-CCTTTTGAGGACTC    4544
B1      TAAACAGGATCTGAATGATCCAGATTTGGGAAGCAGGCCTTTTAAGTACTCTAATATTGCACTTCACCTTTTGAGGACTC    4510
B2      TAAACAGGATCTGAATGATCCAGATTTGGGAAGCAGGCCTTTTAAGTACTCTAATATTGCACTTCACCTTTTGAGGACTC    4545
SIT1    TAAACAGGATCTGGATGATCCAGATTTGGGAAGCAGGCCTTTTAAGTACTCTAATAGTGCACTTCACCTTTTGAGGACTC    4494
BB1     TAAACAGGATCTGAATGATCCAGATTTGGGAAGCAGGCCTTTTAAGTACTCTAATAGTGCACTTCACCTTTTGAGGACTC    4537
BB2     TAAACAGGATCTGAATGATCCAGATTTGGGAAGCAGGCCTTTTAAGTACTCTAATATTGCACTTCACCTTTTGAGGACTC    4546
IMP     TAAATAGGATCTGAATAATCAAGATTTGGGAAGCAGGCCTGGAAAGTACTCTAATATTGCACTTCACCTTTTGAGGACTC    4517
COW     TAAATAGGATCTGAATGATCCAGATTTGGGAAGCAGGCCTGGAAAGTACTCTAATATTGCACTTCACCTTTTGAGGACTC    4492
BUF     TAAATAGGATCTGAATGATCCAGATTTGGGAAGCAGGCCTGGAAAGTACTCTAATATTGCACTTCACCTTTTGAGGACTC    4523
NIL     TAAATAGGATCTGAATGATCCGGATTTGG-AAGCAGGCCTGGAAAGTACTCTAATATTGCACTTCACCTTTTGAGGACTC    4539
```

```
                4570       4580       4590       4600       4610       4620       4630       4640
                   .          .          .          .          .          .          .          .
GE1     TTTAATTTATGTAAATCTAACCATCACCACCACCATCTGTGCTGGATATTGATTGACACGGGTACGTAGTGCATTTCAGT    4626
CE1     TTTAATTTATGTAAATCTAACCATCACCACCACCATCTGTGCTGGATATTGATTGACACGGGTACGTAGTGCATTTCAGT    4618
CE2     TTTAATTTATGTAAATCTAACCATCACCACCACCATCTGTGCTGGATATTGTATGACACGGGTACGTAGTGCATTTCAGT    4625
GK1     TTTAATTTATGTAAATCTAACCATCAACACCACCATCTGTGCTGGATATTGTATGACACGGGTACGTAGTGCATTTCAGT    4618
GK2     TTTAATTTATGTAAATCTAACCATCAACACCACCATCTGTGCTGGATATTGTATGACACGGGTACGTAGTGCATTTCAGT    4626
LK1     TTTAATTTGTGTAAATCTAACCATCAACACCACCATCTGTGCTGGATATTGTATGACACGGGTACGTAGTGCATTTCAGT    4578
CN1     TTTAATTTATGTAAATCTAACCATCAACACCACCATCTGTGCTGGATATTGTATGACACGGGTACGTAGTGCATTTCAGT    4617
CN2     TTTAATTTATGTAAATCTAACCATCAACACCACCATCTGTGCTGGATATTGTATGACACGGGTACGTAGTGCATTTCAGT    4625
MN1     TTTAATTTATGTAAATCTAACCATCAACACCACCATCTGTGCTGGATATTGTATGACACGGGTACGTAGTGCATTTCAGT    4624
MN2     TTTAATTTATGTAAATCTAACCATCAACACCACCATCTGTGCTGGATATTGTATGACACGGGTACGTAGTGCATTTCAGT    4624
B1      TTTAATTTATGTAAATCTAACCATCAACACCACCATCTGTGCTGGATATTGTATGACACGGGTACGTAGTGCATTTCAGT    4590
B2      TTTAATTTATGTAAATCTAACCATCAACACCACCATCTGTGCTGGATATTGTATGACACGGGTACGTAGTGCATTTCAGT    4625
SIT1    TTTAATTTATGTAAATCTAACCATCAACACCACCATCTGTGCTGGATATTGTATGACACGGGTACGTAGTGCATTTCAGT    4574
BB1     TTTAATTTATGTAAATCTAACCATCAACACCACCATCTGTGCTGGATATTGTATGACACGGGTACGTAGTGCATTTCAGT    4617
BB2     TTTAATTTATGTAAATCTAACCATCAACACCACCATCTGTGCTGGATATTGTATGACACGGGTACGTAGTGCATTTCAGT    4626
IMP     TTTAATTTATGTCAATCTAACCATCAACACCACCATCTGTGCTGGATATTGTATGACACGGGTATGTAGTTCATTTCAGT    4597
COW     TTTAATTTATGTAAATCTAACCATCAACACCACCATCTGTGCTGGATATTGTATGACACGGGTACGTAGTGCATTTCAGT    4572
BUF     TTTAATTTATGTAAATCTAACCATCAACACCACCATCTGTGCTGGATATTGTATGACACGGGTATGTAGTGCATTTCAGT    4603
NIL     TTTAATTTATGTAAATCTAACCATCAACACCACCATCTGTGCTGGATATTGTATGACACGGGTACGTAGTGCATTTCAGT    4619

                4650       4660       4670       4680       4690       4700       4710       4720
                   .          .          .          .          .          .          .          .
GE1     TCTTTTAGCTGAAAATTAGATAAACCTAGACCCAGTCCTTTTCTATCCAGAAAGAAAATGAGATAAACCACAGCCTCATT    4706
CE1     TCTTTTAGCTGAAAATTAGATAAACCTAGACCCAGTCCTTTTCTATCCAGAAAGAAAATGAGATAAACCACAACCTCATT    4698
CE2     TCTTTTAGCTGAAAATTAGATAAACCTAGACCCAGTCCTTTTCTATCCAGAAAGAAAATGAGATAAACCACAACCTCATT    4705
GK1     TCTTTTAGCTGAAAATTAGATAAACCTAGACCCAGTCCTTTTCTATCCAGAAAGAAAATGATATAAACCACAACCTCATT    4698
GK2     TCTTTTAGCTGAAAATTAGATAAACCTAGACCCAGTCCTTTTCTATCCAGAAAGAAAATGATATAAACCACAACCTCATT    4706
LK1     TCTTTTAGCTGAAAATTAGATAAACCTAGACCCAGTCCTTTTCTATCCAGAAAGAAAATGAGATAAATCACAACCTCATT    4658
CN1     TCTTTTAGCTGAAAATTAGATAAACCTAGACCCAGTCCTTTTCTATCCAGAAAGAAAATGAGATAAACCACAACCTCATT    4697
CN2     TCTTTTAGCTGAAAATTAGATAAACCTAGACCCAGTCCTTTTCTATCCAGAAAGAAAATGAGATAAACCACAACCTCATT    4705
MN1     TCTTTTAGCTGAAAATTAGATAAACCTAGACCCAGTCCTTTTCTATCCAGAAAGAAAATGAGATAAACCACAACCTCATT    4704
MN2     TCTTTTAGCTGAAAATTAGATAAACCTAGACCCAGTCCTTTTCTATCCAGAAAGAAAATGAGATAAACCACAACCTCATT    4704
B1      TCTTTTAGCTGAAAATTAGATAAACCTAGACCCAGTCCTTTTCTATCCAGAAAGAAAATGAGATAAACCACAACCTCATT    4670
B2      TCTTTTAGCTGAAAATTAGATAAACCTAGACCCAGTCCTTTTCTATCCAGAAAGAAAATGAGATAAACCACAACCTCATT    4705
SIT1    TCTTTTAGCTGAAAATTAGATAAACCTAGACCCAGTCCTTTTCTATCCAGAAAGAAAATGAGATAAACCACAACCTCATT    4654
BB1     TCTTTTAGCTGAAAATTAGATAAACCTAGACCCAGTCCTTTTCTATCCAGAAAGAAAATGAGATAAACCACAACCTCATT    4697
BB2     TCTTTTAGCTGAAAATTAGATAAACCTAGACCCAGTCCTTTTCTATCCAGAAAGAAAATGAGATAAACCACAACCTCATT    4706
IMP     TTTTTTAGCTGAAAATTAGATAAACCTAGACCCAGTCCTTTTCCATCCAGAAAGAAAATGAGATAAACCACAACCTCATT    4677
COW     TCTTTTAGCTGAAAATTAGATAAACCTAGACCCAGTCCTTTTCTATCCAGAAAGAAAATGAGATAAACCACAACCTCATT    4652
BUF     TCTTTTAGCTGAAAATTAGATAAACCTAGACCCAGTCCTTTTCTATCCAGAAAGAAAATGAGATAAACCACAACCTCATT    4683
NIL     TCTTTTAGCTGAAAATTAGATAAACCTAGACCCAGCCCTTTTCTATCCAGAAAGAAAATGAGATAAACCACAACCTCATT    4699

                4730       4740       4750       4760       4770       4780       4790       4800
                   .          .          .          .          .          .          .          .
GE1     TCCAAAATCTAATGCTTGTTGGCTCTGTAGAGGTAGAGTGCACAGGTTACAGTGTATACTCGGCACAGTGGACACAGATA    4786
CE1     TCCAAAATCCAATGCTTGTTGGCTCTGTAGAGGTAGAGTGCACAGGTTACAGTGTATACTCGGCACAGTGGACACAGATA    4778
CE2     TCCAAAATCCAATGCTTGTTGGCTCTGTAGAGGTAGAGTGCACAGGTTACAGTGTATACTCGGCACAGTGGACACAGATA    4785
GK1     TCCAAAATCTAATGCTTGTTGGCTCTGTAGAGGTAGAGTGCACAGGTTACAGTGTATACTCGGCACAGTGGACACAGATA    4778
GK2     TCCAAAATCTAATGCTTGTTGGCTCTGTAGAGGTAGAGTGCACAGGTTACAGTGTATACTCGGCACAGTGGACACAGATA    4786
LK1     TCCAAAATCTAATGCTTGTTGGCTCTGTAGAGGTAGAGTGCACAGGTTACAGTGTATACTCGGCACAGTGGACACAGATA    4738
CN1     TCCAAAATCTAATGCTTGTTGGCTATGTAGAGGTAGAGTGCACAGGTTACAGTGTATACTCAGCACAGTGGACACAGATA    4777
CN2     TCCAAAATCTAATGCTTGTTGGCTATGTAGAGGTAGAGTGCACAGGTTACAGTGTATACTCAGCACAGTGGACACAGATA    4785
MN1     TCCAAAATCTAATGCTTGTTGGCTCTGTAGAGGTAGAGTGCACAGGTTACAGTGTATACTCGGCACAGTGGACACAGATA    4784
MN2     TCCAAAATCTAATGCTTGTTGGCTCTGTAGAGGTAGAGTGCACAGGTTACAGTGTATACTCGGCACAGTGGACACAGATA    4784
B1      TCCAAAATCTAATGCTTGTTGGCTCTGTAGAGGTAGAGTGCACAGATTACAGTGTATACTCGGCACAGTGGACACAGATA    4750
B2      TCCAAAATCTAATGCTTGTTGGCTCTGTAGAGGTAGAGTGCACAGGTTACAGTGTATACTCGGCACAGTGGACACAGATA    4785
SIT1    TCCAAAATCTAATGCTTGTTGGCTCTGTAGAGGTAGAGTGCACAGGTTACAGTGTATACTCGGCACAGTGGACACAGATA    4734
BB1     TCCAAAATCTAATGCTTGTTGGCTCTGTAGAGGTAGAGTGCACAGGTTACAGTGTATACTCGGCACAGTGGACACAGATA    4777
BB2     TCCAAAATCTAATGCTTGTTGGCTCTGTAGAGGTAGAGTGCACAGGTTACAGTGTATACTCGGCACAGTGGACACAGATA    4786
IMP     TCCAAAATCTAATGCTTATTGGCTCTGTAGAGGTAGAGTACACAGGTTACAGTGTATACTCGGGACAGTGGACACAGATA    4757
COW     TCCAAAATCTAATTCTTGTTGGCTCTGTAGAGGTAGAGTGCACAGGTTACAGTGTATACTCAGCACAGTGGACACAGATA    4732
BUF     TCCAAAATCTAATGCTTGTTGGCTCTGTAGAGGTAGAGTGCACAGGTTACAGTGTATACTCAGCACAGTGGACACAGATA    4763
NIL     TCCAAAATCTAATGCTTGCTGGCTCTGTAGAGGTAGAGTGCACAGGTTACAGTGTATACTCAGCACAGTGGATACAGATA    4779
```

```
                4810       4820       4830       4840       4850       4860       4870       4880
                .          .          .          .          .          .          .          .
GE1     ACTGTTTGTAACAGATTT--TTTCAAAATTTATTTAAATGGCATATTGGTCGCATGATGTAAGATGAAAGAGAGGAAGTG    4864
CE1     ACTGTTTGTAACAGATTT--TTTCAAAATTTATTTAAATGGCATATTGGTCGCATGATGTAAGATGAAAGAGAGGAAGTG    4856
CE2     ACTGTTTGTAACAGATTT--TTTCAAAATTTATTTAAATGGCATATTGGTCGCATGATGTAAGATGAAAGAGAGGAAGTG    4863
GK1     ACTGTTTGTAACAGATTT---TTCAAAATTTATTTAAATGGCATATTGGTCGCATGATGTAAGATGAAAGAGAGGAAGTG    4855
GK2     ACTGTTTGTAACAGATTT---TTCAAAATTTATTTAAATGGCATATTGGTCGCATGATGTAAGATGAAAGAGAGGAAGTG    4863
LK1     ACTGTTTGTAACAGATTT--TTTCAAAATTTATTTAAATGGCATATTGGTCGCATGATGTAAGATGAAAGAGAGGAAGTG    4816
CN1     ACTGTTTGTAACAGATTT--TTTCAAAATTTATTTAAATGGCATATTGGTCGCATGATGTAAGATGAAAGAGAGGAAGTG    4855
CN2     ACTGTTTGTAACAGATTT--TTTCAAAATTTATTTAAATGGCATATTGGTCGCATGATGTAAGATGAAAGAGAGGAAGTG    4863
MN1     ACTGTTTGTAACAGATTT--TTTCAAAATTTATTTAAATGGCATATTGGTCACATGATGTAAGATGAAAGAGAGGAAGTG    4862
MN2     ACTGTTTGTAACAGATTT--TTTCAAAATTTATTTAAATGGCATATTGGTCACATGATGTAAGATGAAAGAGAGGAAGTG    4862
B1      ACTGTTTGTAACAGATTT--TTTCAAAATTTATTTAAATGGCATATTGGTCGCATGATGTAAGATGAAAGAGAGGAAGTG    4828
B2      ACTGTTTGTAACAGATTT--TTTCAAAATTTATTTAAATGGCATATTGGTCGCATGATGTAAGATGAAAGAGAGGAAGTG    4863
SIT1    ACTGTTTGTAACAGATTT--TTTCAAAATTTATTTAAATGGCATATTGGTCGCATGATGTAAGATGAAAGAGAGGAAGTG    4812
BB1     ACTGTTTGTAACAGATTT--TTTCAAAATTTATTTAAATGGCATATTGGTCGCATGATGTAAGATGAAAGAGAGGAAGTG    4855
BB2     ACTGTTTGTAACAGATTT--TTTCAAAATTTATTTAAATGGCATATTGGTCGCATGATGTAAGATGAAAGAGAGGAAGTG    4864
IMP     ACTGTTTATAACAGCTTTACTTCCAACATTTATTTAAATGGCATATTGGTCGCATGATGTATACAATGAAAGAAAGGGAGTG    4837
COW     ACTGTTTGTAACAGCTTT--TTCCAAAATTTATTTAAATGGCATATTGGTCGCATGATATAAGATGAAAGAGAGGGAGTG    4810
BUF     ATTGTTTGTAACAGCTTT--TTCCAAAATTTATTTAAATGGCATATTGGTCACATGATATAAGATGAAAGAGAGGGAGTG    4841
NIL     ACTGTTTGTAACAGCTTT--TTCCAAAATGTATTTAAATGGCATATTAGTTGCATGATATAAGATGAAAGAGAGGGAGTG    4857

                4890       4900       4910       4920       4930       4940       4950       4960
                .          .          .          .          .          .          .          .
GE1     TCACTTTTATCTCTACGGGATTTAGTGTGGATATATTAAGTTGCCATTGGGAAATGGAATTAAAGAATCCTCCTTGAGTC    4944
CE1     TCACTTTTATCTCTACGGGATTTAGTGTGGATATATTAAGTTGCCATTGGGAAATGGAATTAAAGAATCCTCCTTGAGTC    4936
CE2     TCACTTTTATCTCTACGGGATTTAGTGTGGATATATTAAGTTGCCATTGGGAAATGGAATTAAAGAATCCTCCTTGAGTC    4943
GK1     TCACTTTTATCTCTATGGGATTTAGTGTGGATATATTAAGTTGCCATTGGGAAATGGAATTAAAGAATCCTCCTTGAGTC    4935
GK2     TCACTTTTATCTCTATGGGATTTAGTGTGGATATATTAAGTTGCCATTGGGAAATGGAATTAAAGAATCCTCCTTGAGTC    4943
LK1     TCACTTTTATCTCTACGGGATTTAGTGTGGATATATTAAGTTGCCATTGGGAAATGGAATTAAAGAATCCTCCTTGAGTC    4896
CN1     TCACTTTTATCTCTACGGGATTTAGTGTGGATATATTAAGTTGCCATTGGGAAATGGAATTAAAGAATCCTCCTTGAGCC    4935
CN2     TCACTTTTATCTCTACGGGATTTAGTGTGGATATATTAAGTTGCCATTGGGAAATGGAATTAAAGAATCCTCCTTGAGCC    4943
MN1     TCACTTTTATCTCTACAGGATTTAGTGTGGATATATGAAGTTGCCATTGGGAAATGGAATTAAAGAATCCTCCTTGAGTC    4942
MN2     TCACTTTTATCTCTACAGGATTTAGTGTGGATATATGAAGTTGCCATTGGGAAATGGAATTAAAGAATCCTCCTTGAGTC    4942
B1      TCACTTTTATCTCTACGGGATTTAGTGTGGATATATTAAGTTGCCATTGGGAAATGGAATTAAAGAATCCTCCTTGAGTC    4908
B2      TCACTTTTATCTCTACGGGATTTAGTGTGGATATATTAAGTTGCCATTGGGAAATGGAATTAAAGAATCCTCCTTGAGTC    4943
SIT1    TCACTTTTATCTCTACGGGATTTAGTGTGGATATATTAAGTTGCCATTGGGAAATGGAATTAAAGAATCCTCCTTGAGTC    4892
BB1     TCACTTTTATCTCTACGGGATTTAGTGTGGATATATTAAGTTGCCATTGGGAAATGGAATTAAAGAATCCTCCTTGAGTC    4935
BB2     TCACTTTTATCTCTACGGGATTTAGTGTGGATATATTAAGTTGCCATTGGGAAATGGAATTAAAGAATCCTCCTTGAGTC    4944
IMP     TCACTTTTATCTCTATGGGATTTAATGTGGATATATTGAGTTGGCATTGGGAAATGGAACTAAAGAATCCTCCTCTAGTC    4917
COW     TCACTTTTATCTCAACAGGGTTTAGTGTGGATATATTAAGTTGGCATTGGGAAATGGAACTAAAGAATCCTCCTCGAGTC    4890
BUF     TCACTTTTATCTCAATGGGATTTAGTGTGGATATATTAAGTTGGCATTGGGAAATGGAACTAAAGAATCCTCCTCTAGTC    4921
NIL     TCACTTTTATCTCAATGGGATTTAGTGTGGATATATTAAGTTGGCATTGGGAAACGGAACTAAAGAATCCTCCTTGAGTC    4937

                4970       4980       4990       5000       5010       5020       5030       5040
                .          .          .          .          .          .          .          .
GE1     CTATTTATGACAAAATAATAT--GTTAATTAATTTTTATATGTTTCCATTACTATATATTTTTCTAAACTCCTATCGCATT    5022
CE1     CTATTTATGACAAAATAATAT--GTTAATTAATTTTTATATGTTTCCATTACTATATATTTTTCTAAACTCCTATCGCATT    5014
CE2     CTATTTATGACAAAATAATAT--GTTAATTAATTTTTATATGTTTCCATTACTATATATTTTTCTAAACTCCTATCGCATT    5021
GK1     CTATTTATGACAAAATAATAT--GTTAATTAATTTTTATATGTTTCCATTACTATATATTTTTCTAAACTCCTATCACATT    5013
GK2     CTATTTATGACAAAATAATAT--GTTAATTAATTTTTATATGTTTCCATTACTATATATTTTTCTAAACTCCTATCACATT    5021
LK1     CTATTTATGACAAAATAATAT--GTTAATTAATTTTTATATGTTTCCATTACTATATATTTTTCTAAACTCCTATCACATT    4974
CN1     CTATTTATGACAAAATAATAT--GTTAATTAATTTTTATATGTTTCCATTACTATATATTTTTCTAAACTCCTATCGCATT    5013
CN2     CTATTTATGACAAAATAATAT--GTTAATTAATTTTTATATGTTTCCATTACTATATATTTTTCTAAACTCCTATCGCATT    5021
MN1     CTATTTATGACAAAATAATAT--GTTAATTAATTTTTATATGTTTCCATTACTATATATTTTTCTAAACTCCTATCACATT    5020
MN2     CTATTTATGACAAAATAATAT--GTTAATTAATTTTTATATGTTTCCATTACTATATATTTTTCTAAACTCCTATCACATT    5020
B1      CTATTTATGACAAAATAATAT--GTTAATTAATTTTTATATGTTTCCATTACTATATATTTTTCTAAACTCCTATCGCATT    4986
B2      CTATTTATGACAAAATAATAT--GTTAATTAATTTTTATATGTTTCCATTACTATATATTTTTCTAAACTCCTATCGCATT    5021
SIT1    CTATTTATGACAAAATAATAT--GTTAATTAATTTTTATATGTTTCCATTACTATATATTTTTCTAAACTCCTATCGCATT    4970
BB1     CTATTTATGACAAAATAATAT--GTTAATTAATTTTTATATGTTTCCATTACTATATATTTTTCTAAACTCCTATCGCATT    5013
BB2     CTATTTATGACAAAATAATAT--GTTAATTAATTTTTATATGTTTCCATTACTATATATTTTTCTAAACTCCTATCGCATT    5022
IMP     CTATTTATGACAAAATAATATAAGTTAATT----TTTATCTGTTTCCATTACTATATATTTTTCTAAAGTCCTATCGCATT    4993
COW     CTATTTATGACAAAATAAT----GTTAATTAATTTTTATCTGTTTCCATTACTATATATTTTTCTAAAGTCCTATCGCATT    4966
BUF     CTATTTATGACAAAATAATAT--GTTAATTAATTTTTATCTGTTTCCATTACTATATATTTTTCTAAAGTCCTATTGCATT    4999
NIL     CTATTTATGACAAAATAATAT--GTTAATTAATTTTTATCTGTTTCCATTACTATATACTTTCTAAAGTCCTATTGCATT    5015
```

```
                5050       5060       5070       5080       5090       5100       5110       5120
                .          .          .          .          .          .          .          .
GE1     ATGCTCCTTTTAAATTCTTTCCTCAGGATGTCAATGGCAAGCTGTTTCTTCCCAAATATGCCCTGTCTCAGGATGTCTGT    5102
CE1     ATGCTCCTTTTAAATTCTTTCCTCAGGATGTCAACGGCAAGCTGTTTCTTCCCAAATATGCCCTGTCTCAGGATGTCTGT    5094
CE2     ATGCTCCTTTTAAATTCTTTCCTCAGGATGTCAACGGCAAGCTGTTTCTTCCCAAATATGCCCTGTCTCAGGATGTCTGT    5101
GK1     ATGCTCCTTTTAAATTCTTTCCTCAGGATGTCAATGGCAAGCAGTTTCTTCCCAAATATGCCCTGTCTCAGGATGTCTGT    5093
GK2     ATGCTCCTTTTAAATTCTTTCCTCAGGATGTCAATGGCAAGCAGTTTCTTCCCAAATATGCCCTGTCTCAGGATGTCTGT    5101
LK1     ATGCTCCTTTTAAATTCTTTCCTCAGGATGTCAACGCCAAGCTGTTTCTTCCCAAATATGCCCTGTCTCAGGATGTCTGT    5054
CN1     ATGCTCCTTTTAAATTCTTTCCTCAGGATGTCAATGGCAAGCTGTTTCTTCCCAAATATGCCCTGTCTCAGGATGTCTGT    5093
CN2     ATGCTCCTTTTAAATTCTTTCCTCAGGATGTCAATGGCAAGCTGTTTCTTCCCAAATATGCCCTGTCTCAGGATGTCTGT    5101
MN1     ATGCTCCTTTTAAATTCTTTCCTCAGGATGTCAATGGCAAGCAGTTTCTTCCCAAATATGCCCTGTCTCAGGATGTCTGT    5100
MN2     ATGCTCCTTTTAAATTCTTTCCTCAGGATGTCAATGGCAAGCAGTTTCTTCCCAAATATGCCCTGTCTCAGGATGTCTGT    5100
B1      ATGCTCCTTTTAAATTCTTTCCTCAGGATGTCAATGGCAAGCTGTTTCTTCCCAAATATGCCCTGTCTCAGGATGTCTGT    5066
B2      ATGCTCCTTTTAAATTCTTTCCTCAGGATGTCAATGGCAAGCTGTTTCTTCCCAAATATGCCCTGTCTCAGGATGTCTGT    5101
SIT1    ATGCTCCTTTTAAATTCTTTCCTCAGGATGTCAATGGCAAGCTGTTTCTTCCCAAATATGCCCTGTCTCAGGATGTCTGT    5050
BB1     ATGCTCCTTTTAAATTCTTTCCTCAGGATGTCAATGGCAAGCTGTTTCTTCCCAAATATGCCCTGTCTCAGGATGTCTGT    5093
BB2     ATGCTCCTTTTAAATTCTTTCCTCAGGATGTCAATGGCAAGCTGTTTCTTCCCAAATATGCCCTGTCTCAGGATGTCTGT    5102
IMP     ATGCTCCTTTTAAATTCTTTCCTCAGGATGTCAACGGCAAGCTGTTTCTTCCCAAATATGCCCTCTCTCAGGATGTCTGT    5073
COW     ATGCTCCTTTTAAATTCTTTCCTCAGGATGTCAACGGCAAGCTGTTTCTTCCCAAATATGCCCTGTCTCAGGATGTGTGT    5046
BUF     ATGCTCCTTTTAAATTCTTTCCTCAGGATGTCAACGGCAAGCTGTTTCTTCCCAAATATGCCCTGTCTCAGGATGTCTGT    5079
NIL     ATGCTCCTTTTAAATTCTTTCCTCAGGATGTCAACGGCAAGCTGTTTCTTCCCAAATATGCCCTGTCTCAGGATGTCTGT    5095

                5130       5140       5150       5160       5170       5180       5190       5200
                .          .          .          .          .          .          .          .
GE1     ACATACAGAGACTTCATGTACAAGACTGCAGAAATACCAGGATGCCCACGCCATGTTACTCCTTATTTCTCCTACCCGGT    5182
CE1     ACATACAGAGACTTCATGTACAAGACTGCAGAAATACCAGGATGCCCACGCCATGTTACTCCTTATTTCTCCTACCCGGT    5174
CE2     ACATACAGAGACTTCATGTACAAGACTGCAGAAATACCAGGATGCCCACGCCATGTTACTCCTTATTTCTCCTACCCGGT    5181
GK1     ACATACAGAGACTTCATGTACAAGACTGCAGAAATACCAGGATGCCCACGCCATGTTACTCCTTATTTCTCCTACCCGGT    5173
GK2     ACATACAGAGACTTCATGTACAAGACTGCAGAAATACCAGGATGCCCACGCCATGTTACTCCTTATTTCTCCTACCCGGT    5181
LK1     ACATACAGAGACTTCATGTACAAGACTGCAGAAATACCAGGATGCCCACGCCATGTTACTCCTTATTTCTCCTACCCGGT    5134
CN1     ACATACAGAGACTTCATGTACAAGACTGCAGAAATACCAGGATGCCCACGCCATGTTACTCCTTATTTCTCCTACCCGGT    5173
CN2     ACATACAGAGACTTCATGTACAAGACTGCAGAAATACCAGGATGCCCACGCCATGTTACTCCTTATTTCTCCTACCCGGT    5181
MN1     ACATACAGAGACTTCATGTACAAGACTGCAGAAATACCAGGATGCCCACGCCATGTTACTCCTTATTTCTCCTACCCGGG    5180
MN2     ACATACAGAGACTTCATGTACAAGACTGCAGAAATACCAGGATGCCCACGCCATGTTACTCCTTATTTCTCCTACCCGGG    5180
B1      ACATACAGAGACTTCATGTACAAGACTGCAGAAATACCAGGATGCCCACGCCATGTTACTCCTTATTTCTCCTACCCGGT    5146
B2      ACATACAGAGACTTCATGTACAAGACTGCAGAAATACCAGGATGCCCACGCCATGTTACTCCTTATTTCTCCTACCCGGT    5181
SIT1    ACATACAGAGACTTCATGTACAAGACTGCAGAAATACCAGGATGCCCACGCCATGTTACTCCTTATTTCTCCTACCCGGT    5130
BB1     ACATACAGAGACTTCATGTACAAGACTGCAGAAATACCAGGATGCCCACGCCATGTTACTCCTTATTTCTCCTACCCAGT    5173
BB2     ACATACAGAGACTTCATGTACAAGACTGCAGAAATACCAGGATGCCCACGCCATGTTACTCCTTATTTCTCCTACCCAGT    5182
IMP     ACATACAGAGACTTCATGTACAAGACTGCAGAAATACCAGGATGCCCACGCCATGTTACTCCTTATTTCTCCTACCCGGT    5153
COW     ACATACAGAGACTTCATGTACAAGACTGCAGAAATACCAGGATGCCCACGCCATGTTACTCCTTATTTCTCCTACCCGGT    5126
BUF     ACATACAGAGATTTCATGTACAAGACTGCAGAAATACCAGGATGCCCACGCCATGTTACTCCTTATTTCTCCTACCCGGT    5159
NIL     ACATATAGAGACTTCATGTACAAGACTGCAGAAATACCAGGATGCCCACGCCATGTTACTCCTTATTTCTCCTACCCGGT    5175

                5210       5220       5230       5240       5250       5260       5270       5280
                .          .          .          .          .          .          .          .
GE1     GGCTATAAGCTGTAAGTGTGGCAAGTGTAATACTGACTATAGTAAGCATGCCCTGGTTGAAGCAGACATTGGCATCCAGG    5262
CE1     AGCTATAAGCTGTAAGTGTGGCAAGTGTAATACTGACTATAGTAAGCATGCCCTGGTTGAAGCAGACATTGGCATCCAGG    5254
CE2     AGCTATAAGCTGTAAGTGTGGCAAGTGTAATACTGACTATAGTAAGCATGCCCTGGTTGAAGCAGACATTGGCATCCAGG    5261
GK1     GGCTATAAGCTGTAAGTGTGGCAAGTGTAATACTGACTATAGTAAGCATGCCCTGGTTGAAGCAGACATTGGCATCCAGG    5253
GK2     GGCTATAAGCTGTAAGTGTGGCAAGTGTAATACTGACTATAGTAAGCATGCCCTGGTTGAAGCAGACATTGGCATCCAGG    5261
LK1     AGCTATAAGCTGTAAGTGTGGCAAGTGTAATACTGACTATAGTAAGCATGCCCTGGTTGAAGCAGACATTGGCATCCAGG    5214
CN1     AGCTATAAGCTGTAAGTGTGGCAAGTGTAATACTGACTATAGTAAGCATGCCCTGGTTGAAGCAGACATTGGCATCCAGG    5253
CN2     AGCTATAAGCTGTAAGTGTGGCAAGTGTAATACTGACTATAGTAAGCATGCCCTGGTTGAAGCAGACATTGGCATCCAGG    5261
MN1     GGCTATAAGCTGTAAGTGTGGCAAGTGTAATACTGACTATAGTAAGCATGCCCTGGTTGAAGCAGACATTGGCATCCAGG    5260
MN2     GGCTATAAGCTGTAAGTGTGGCAAGTGTAATACTGACTATAGTAAGCATGCCCTGGTTGAAGCAGACATTGGCATCCAGG    5260
B1      GGCTATAAGCTGTAAGTGTGGCAAGTGTAATACTGACTATAGTAAGCATGCCCTGGTTGAAGCAGACATTGGCATCCAGG    5226
B2      GGCTATAAGCTGTAAGTGTGGCAAGTGTAATACTGACTATAGTAAGCATGCCCTGGTTGAAGCAGACATTGGCATCCAGG    5261
SIT1    GGCTATAAGCTGTAAGTGTGGCAAGTGTAATACTGACTATAGTAAGCATGCCCTGGTTGAAGCAGACATTGGCATCCAGG    5210
BB1     GGCTATAAGCTGTAAGTGTGGCAAGTGTAATACTGACTATAGTAAGCATGCCCTGGTTGAAGCAGACATTGGCATCCAGG    5253
BB2     GGCTATAAGCTGTAAGTGTGGCAAGTGTAATACTGACTATAGTAAGCATGCCCTGGTTGAAGCGGACATTGGCATCCAGG    5262
IMP     AGCTATAAGCTGTAAGTGTGGCAAGTGTAATACTGACTATAGTAAGCACGCCCTGGTTGAAGCAGACATTGGCATCCAGG    5233
COW     AGCTATAAGCTGTAAGTGTGGCAAGTGTAATACTGACTATAGTAAGCATGCCCTGGTTGAAGCAGACATTGGCATCCAGG    5206
BUF     AGCTATAAGCTGTAAGTGTGGCAAGTGTAATACTGACTATAGTAAGCATGCCCTGGTTGAAGCAGACATTGGCATCCAGG    5239
NIL     AGCTATAAGCTGTAAGTGTGGCAAGTGTAATACTGACTATAGTAAGCATGCCCTGGTTGAAGCAGACATTGGCATCCAGG    5255
```

```
                5290       5300       5310       5320       5330       5340       5350       5360
                .          .          .          .          .          .          .          .
GE1     CAGAGCGGGTGAGAGGTGTCAACGCTTCTGCCCAGAAGTTCGCAACCGATGGGGAAGGTAAGGATGGACCGACCCCGACG     5342
CE1     CAGAGCGGGTGAGAGGTGTCAACGCTTCTGCCCAGAAGTTCGCAACCGATGGGGAAGGTAAGGATGGACCGACCCCGACG     5334
CE2     CAGAGCGGGTGAGAGGTGTCAACGCTTCTGCCCAGAAGTTCGCAACCGATGGGGAAGGTAAGGATGGACCGACCCCGACG     5341
GK1     CAGAGCGGGTGAGAGGTGTCAACGCTTCTGCCCAGAAGTTCGCAACCGATGGGGAAGGTAAGGATGGACCGACCCCGACG     5333
GK2     CAGAGCGGGTGAGAGGTGTCAACGCTTCTGCCCAGAAGTTCGCAACCGATGGGGAAGGTAAGGATGGACCGACCCCGACG     5341
LK1     CAGAGCGGGTGAGAGGTGTCAACGCTTCTGCCCAGAAGTTCGCAACCGATGGGGAAGGTAAGGATGGACCGACCCCGACG     5294
CN1     CAGAGCGGGTGAGAGGTGTCAACGCTTCTGCCCAGAAGTTCGCAACCGATGGGGAAGGTAAGGATGGACCGACCCCAACA     5333
CN2     CAGAGCGGGTGAGAGGTGTCAACGCTTCTGCCCAGAAGTTCGCAACCGATGGGGAAGGTAAGGATGGACCGACCCCAACA     5341
MN1     CAGAGCGGCTGAGAGGTGTCAACGCTTCTGCCCAGAAGTTCGCAACCGATGGGGAAGGTAAGGATGGACCGACCCCGACG     5340
MN2     CAGAGCGGCTGAGAGGTGTCAACGCTTCTGCCCAGAAGTTCGCAACCGATGGGGAAGGTAAGGATGGACCGACCCCGACG     5340
B1      CAGAGCGGGTGAGAGGTGTCAACGCTTCTGCCCAGAAGTTCGCAACCGATGGGGAAGGTAAGGATGGACCGACCCCGACG     5306
B2      CAGAGCGGGTGAGAGGTGTCAACGCTTCTGCCCAGAAGTTCGCAACCGATGGGGAAGGTAAGGATGGACCGACCCCGACG     5341
SIT1    CAGAGCGGGTGAGAGGTGTCAACGCTTCTGCCCAGAAGTTCGCAACCGATGGGGAAGGTAAGGATGGACCGACCCCGACG     5290
BB1     CAGAGCGGGTGAGAGGTGTCAACGCTTCTGCCCAGAAGTTCGCAACCGATGGGGAAGGTAAGGATGGACCGACCCCGACG     5333
BB2     CAGAGCGGGTGAGAGGTGTCAACGCTTCTGCCCAGAAGTTCGCAACCGATGGGGAAGGTAAGGATGGACCGACCCCGACG     5342
IMP     CAGAGCGGGTGAGAGGTGTCAACGCTTCTGCCCAGAAGTTCGCAACCGATGGGGAAGGTAAGGATGGACCGACCCTGACG     5313
COW     CAGAGCGGGTGAGAGGTGTCAACGCTTCTGCCCAGAAGTTTGCAACCGACGGGGAAGGTAAGGATGGACCGACCCCCACA     5286
BUF     CAGAGCGGGTGAGAGGTGTCAACGCTTCTGCCCAGAAGTTCGCAACCGACGGGGAAGGTAAGGATGGACCAACCCCAACG     5319
NIL     CAGAGCGGGTGAGAGGTGTCAACGCTTCTGCCCAGAAGTTCGCAACCGACGGGGAAGGTAAGGATGGACTGACCCCAACG     5335

                5370       5380       5390       5400       5410       5420       5430       5440
                .          .          .          .          .          .          .          .
GE1     GTACCTCCTCATCACCAGAGGCCCGTCTCTAGAGACTTGGATGAGCTATGTGTGGTAGTGCAGCCTTGAAAGGTACCAGC     5422
CE1     GTACTTCCTCATCACCAGAGGCCCGTCTCTAGAGACTTGGATGAGCTATGTGTGGTAGTGCAGCCTTGAAAGGTACCAGC     5414
CE2     GTACTTCCTCATCACCAGAGGCCCGTCTCTAGAGACTTGGATGAGCTATGTGTGGTAGTGCAGCCTTGAAAGGTACCAGC     5421
GK1     GTACCTCCTCATCACCAGAGGCCCGTCTCTAGAGACTTGGATGAGCTATGTGTGGTAGTGCAGCCTTGAAAGGTACCAGC     5413
GK2     GTACCTCCTCATCACCAGAGGCCCGTCTCTAGAGACTTGGATGAGCTATGTGTGGTAGTGCAGCCTTGAAAGGTACCAGC     5421
LK1     GTACCTCCTCATCACCAGAGGCCCGTCTCTAGAGACTTGGATGAGCTATGTGTGGTAGTGCAGCCTTGAAAGGTACCAGC     5374
CN1     GTACCTCCTCATCACCAGAGGCTCGTCTCTAGAGACTTGGATGAGCTATGTGTGGTAGTGCAGCCTTGAAAGGTACCAGC     5413
CN2     GTACCTCCTCATCACCAGAGGCTCGTCTCTAGAGACTTGGATGAGCTATGTGTGGTAGTGCAGCCTTGAAAGGTACCAGC     5421
MN1     GTACCTCCTCATCACCAGAGGCCCGTCTCTAGAGACTTGGATGAGCTATGTGTGGTAGTGCAGCCTTGAAAGGTACCAGC     5420
MN2     GTACCTCCTCATCACCAGAGGCCCGTCTCTAGAGACTTGGATGAGCTATGTGTGGTAGTGCAGCCTTGAAAGGTACCAGC     5420
B1      GTACCTCCTCATCACCAGAGGCCCGTCTCTAGAGACTTGGATGAACTATGTGTGGTAGTGCAGCCTTGAAAGGTACCAGC     5386
B2      GTACCTCCTCATCACCAGAGGCCCGTCTCTAGAGACTTGGATGAACTATGTGTGGTAGTGCAGCCTTGAAAGGTACCAGC     5421
SIT1    GTACCTCCTCATCACCAGAGGCCCGTCTCTAGAGACTTGGATGAACTATGTGTGGTAGTGCAGCCTTGAAAGGTACCAGC     5370
BB1     GTACCTCCTCATCACCAGAGACCCGTCTCTAGAGACTTGGATGAGCTATGTGTGGTAGTGCAGCCTTGAAAGGTACCAGC     5413
BB2     GTACCTCCTCATCACCAGAGACCCGTCTCTAGAGACTTGGATGAGCTATGTGTGGTAGTGCAGCCTTGAAAGGTACCAGC     5422
IMP     GTACCTCCTCATCACCAGAGGCCCGTCTCTAGAGACTTGGATGAGCTGTGTGTGGTAGTGCAGCCTTGAAAGGTACCAGC     5393
COW     GTACCTCCTCATCACCAGAGGCCCATCTCTAGAGACTTGGATGAGCTGTGTGTGGTAGTGCAGCCTTGAAAGGTACCAGC     5366
BUF     GTACCTCCTCATCATCAGAGGCCCGTCTCTAGAGACTTGGATGAGCTATGTGTGGTAGTGCAGCCTTGAAAGGTACCAGC     5399
NIL     GTACCTCCTCATCACCAGAGGCCCGTCTCTAGAGACTTGGATAACCTGTGTGTGGTAGTGCAACCTTGAAAGGTACCAGC     5415

                5450       5460       5470       5480       5490       5500       5510       5520
                .          .          .          .          .          .          .          .
GE1     CATTTGTGGGTTCCCAGT--------------------------------AGGGTGGCCTTCTGAACACTTTTAACTCGTT     5471
CE1     CATTTGTGGGTTCCCAGT--------------------------------AGGGTGGCCTTCTGAACACTTTTAACTCGTT     5463
CE2     CATTTGTGGGTTCCCAGT--------------------------------AGGGTGGCCTTCTGAACACTTTTAACTCGTT     5470
GK1     CATTTGTGGGTTCCCAGT--------------------------------AGGGTGGCCTTCTGAACACTTTTAACTCGTT     5462
GK2     CATTTGTGGGTTCCCAGT--------------------------------AGGGTGGCCTTCTGAACACTTTTAACTCGTT     5470
LK1     CATTTGTGGGTTCCCAGTGGGCCGTGTTGGGGATATAACATGATGCTGTAGGGTGGCCTTCCAAACACTTTTAACTTGTT     5454
CN1     CATTTGTGGGTTCCCAGTGGGCCGTGTTGGGGATATAACATGATGCTATAGGGTGGCCTTCTGAACACTTTAAACTCGTT     5493
CN2     CATTTGTGGGTTCCCAGTGGGCCGTGTTGGGGATATAACATGATGCTATAGGGTGGCCTTCTGAACACTTTAAACTCGTT     5501
MN1     CACTTGTGGGTTCCCAGT--------------------------------AGGGTGGCCTTCTGAACACTTTTAACTCGTT     5469
MN2     CACTTGTGGGTTCCCAGT--------------------------------AGGGTGGCCTTCTGAACACTTTTAACTCGTT     5469
B1      CATTTGTGGGTTCCCAGT--------------------------------AGGGTGGCCTTCTGAACACTTTTAACTCGTT     5435
B2      CATTTGTGGGTTCCCAGT--------------------------------AGGGTGGCCTTCTGAACACTTTTAACTCGTT     5470
SIT1    CATTTGTGGGTTCCCAGT--------------------------------AGGGTGGCCTTTTGAACACTTTTAACTAGTT     5419
BB1     CATTTGTGGGTTCCCAGT--------------------------------AGGGTGGCCTTCTGAACACTTTTAACTCGTT     5462
BB2     CATTTGTGGGTTCCCAGT--------------------------------AGGGTGGCCTTCTGAACACTTTTAACTCGTT     5471
IMP     CATTTGTGGGTTCCCAGTGGGCCATGTTGGGGATTTAACAAGATGCTGTAGGGTGGCCTTCTGAACATTTTTAACTGCTT     5473
COW     CACTTGTGGGTTCCCAGTGGGCCGTGTTGGGGATATAACATGATGCTGTAGGGTGGCCTTCTGAACACTTTTAACTCCTT     5446
BUF     CATTTGTGGGTTCCCAGTGGGCCATGTTGGGGATATAACATGATGCTGTAGGGTGGCCTTCTGAACACTTTTAACTCCTT     5479
NIL     CGTTTGTGGGTTCCCAGTTGGCCATGTTGGGGATATAACATGATGCTGTAGGGTGGCCTTCTGAACACTTTTAACTCCTT     5495
```

```
                5530       5540       5550       5560       5570       5580       5590       5600
                .          .          .          .          .          .          .          .
GE1     CTTACTTTTCTGAAATGGACCTATCTTTCCCTTCAGTACCGAAGTGCTACTTGCTTATGTTGGGC-AGTGAACTTTGCCA    5550
CE1     CTTTCTTTTCTGAAATGGACCTATCTTTCCCTTCAGTACCGAAGTGCTACTTGATTATGTTGGGC-AGTGAACTTTGCCA    5542
CE2     CTTTCTTTTCTGAAATGGACCTATCTTTCCCTTCAGTACCGAAGTGCTACTTGATTATGTTGGGC-AGTGAACTTTGCCA    5549
GK1     CTTACTTTTCTGAAATGGACCTATCTTTCCCTTCAGTACCGAAGTGCTACTTGATTATGTTGGGC-AGTGAACTTTGCCA    5541
GK2     CTTACTTTTCTGAAATGGACCTATCTTTCCCTTCAGTACCGAAGTGCTACTTGATTATGTTGGGC-AGTGAACTTTGCCA    5549
LK1     CTTACTTTTCTGAAATGGACCTATCTTTCCCTTCAGTACCGAAATGCTACTTGCTTATGTTGGGC-AGTGAACTTTGCCA    5533
CN1     CTTACTTTTCTGAAATGGACCTATCTTTCCCTTCAGTACCGAAGTGCTACTTGCTTATGTTGGGC-AGTGAACTTTGCCA    5572
CN2     CTTACTTTTCTGAAATGGACCTATCTTTCCCTTCAGTACCGAAGTGCTACTTGCTTATGTTGGGC-AGTGAACTTTGCCA    5580
MN1     CTTACTTTTCTGAAATGGATCTATCTTTCCCTTTAGTACCGAAGTGCTACTTGCTTATGTTGGGC-AGTGAACTTTGCCA    5548
MN2     CTTACTTTTCTGAAATGGATCTATCTTTCCCTTTAGTACCGAAGTGCTACTTGCTTATGTTGGGC-AGTGAACTTTGCCA    5548
B1      CTTTCTTTTCTGAAATGGACCTATCTTTCCCTTTAGTACCGAAGTGCTACTTGCTTATGTTGGGC-AGTGAACTTTGCCA    5514
B2      CTTTCTTTTCTGAAATGGACCTATCTTTCCCTTTAGTACCGAAGTGCTACTTGCTTATGTTGGGC-AGTGAACTTTGCCA    5549
SIT1    CTTACTTTTCTGAAATGGACCTATCTTTCCCTTTAGTACCGAAGTGCTACTTGCTTATGTTGGGC-AGTGAACTTTGCCA    5498
BB1     CTTACTTTTCTGAAATGGACCTATCTTTCCCTTTAGTACTGAAGTGCTACTTGCTTATGTTGGGC-AGTGAACTTTGCCA    5541
BB2     CTTACTTTTCTGAAATGGACCTATCTTTCCCTTTAGTACTGAAGTGCTACTTGCTTATGTTGGGC-AGTGAACTTTGCCA    5550
IMP     ---ACTTTTCTGAAGTGGACCTATCTTTTCCTTCAGTACCTAAGTGCTACTTGCTTATGTTGGGCAAGTGAACTTTGCCA    5550
COW     CTTACTTTTCTGAAATGGACCTATCTTTCCCTTCAGTACCCAAGTGCTACTTGCTTATGTTGGGCAAGTGAACTTTGCCA    5526
BUF     T---CTTTTCTGAAATGGACCTATCTTTCCCTTCAGTACCCAAGTGCTACTTGCTTACGTTGGGCAAGTGAACTTTGCCA    5556
NIL     ----CTTTTCTGAAATGAACCTATCTTTCCCTTCAGTACCCAAGTGCTACTTGCTTATGTTGGGCAAGTGAACTTTGCTA    5571

                5610       5620       5630       5640       5650       5660       5670       5680
                .          .          .          .          .          .          .          .
GE1     GACCACACTCTGCACACCCACTGCCTGGCATGGCTCTCTTGGCTTTCACACCTGGTCGTGCAGATATTTGTTCTAAGTTT    5630
CE1     GACCACACTCTGCACACCCACTGCCTGGCATGGCTCTCTTGGCTTTCACACCTGGTCGTGCAGATATTTGTTCTAAGTTT    5622
CE2     GACCACACTCTGCACACCCACTGCCTGGCATGGCTCTCTTGGCTTTCACACCTGGTCGTGCAGATATTTGTTCTAAGTTT    5629
GK1     GACCACACTCTGCACACCCACTGCCTGGCATGGCTCTCTTGGCTTTCACACCTGGTCATGCAGATATTTGTTCTAAGTTT    5621
GK2     GACCACACTCTGCACACCCACTGCCTGGCATGGCTCTCTTGGCTTTCACACCTGGTCATGCAGATATTTGTTCTAAGTTT    5629
LK1     GACCACACTCTGCACACTCACTGCTTGGCATGGCTCTCTTAGCTTTCACACCTGGTCGTGCACATATTTGTTCTAAGTTT    5613
CN1     GACCACACTCTGCACACCCACTGCCTGGCATGGCTCTCTTAGCTTTCACACCTGGTCGTGCAGATATTTGTTCTAAGTTT    5652
CN2     GACCACACTCTGCACACCCACTGCCTGGCATGGCTCTCTTAGCTTTCACACCTGGTCGTGCAGATATTTGTTCTAAGTTT    5660
MN1     GAGCACACTCTGCACACCCACTGCCTGGCATGGCTCTCTTGGCTTTCACACCTGGTCGTGCAGATATTTTTTCTAAGTTT    5628
MN2     GAGCACACTCTGCACACCCACTGCCTGGCATGGCTCTCTTGGCTTTCACACCTGGTCGTGCAGATATTTTTTCTAAGTTT    5628
B1      GAGCACACTCTGCACACCCACTGCCTGGCATGGCTCTCTTGGCTTTCACACCTGGTCGTGCAGATATTTTTTCTAAGTTT    5594
B2      GAGCACACTCTGCACACCCACTGCCTGGCATGGCTCTCTTGGCTTTCACACCTGGTCGTGCAGATATTTTTTCTAAGTTT    5629
SIT1    GAGCACACTCTGCACACCCACTGCCTGGCATGGCTCTCTTGGCTTTCACACCTGGTCGTGCAGATATTTTTTCTAAGTTT    5578
BB1     GAGCACACTCTGCACACCCACTGCCTGGCATGGCTCTCTTGGCTTTCACACCTGGTCGTGCAGATATTTTTTCTAAGTTT    5621
BB2     GAGCACACTCTGCACACCCACTGCCTGGCATGGCTCTCTTGGCTTTCACACCTGGTCGTGCAGATATTTTTTCTAAGTTT    5630
IMP     GACTACACTCTGCACACCCACTGCCTGGCGTGGCTCTCTTGGCTTTCACACCTGGTCGTGCAGATATTTGTTCCACGTTT    5630
COW     GACCACACTCTGCACACCCACTGCCTGGCATGGCTCTCTTGGCTTTCACACCTGGTCGTGCAGATATTTGTTCTAAGTTT    5606
BUF     GACCACACTCTGCACACCCACTGCCTGGCATGGCTCTCTTGGCTTTCACACCTGGCCATGCAGATATTTGTTCTAAGTGT    5636
NIL     GCCCACACTCTGCACACCCACTGCCTGGCATGGCTCTCTTGGCTTTCACACCTGGTCGTGCAGATATTTGTTCTAAGTTT    5651

                5690       5700       5710       5720       5730       5740       5750       5760
                .          .          .          .          .          .          .          .
GE1     GTTTGGAAAGGATTGGTTTTGTGCATATTCAGAAACATTTAAATGCAGCTGCCTTCTTGCTTATAGTTTGCCTCGTTTTT    5710
CE1     GTTTGGAAAGGATTGGTTTTGTGCATATTCAGAAACATTTAAATGCAGCTGCCTTCTTGCTTATAGTTTGCCTCGTTTTT    5702
CE2     GTTTGGAAAGGATTGGTTTTGTGCATATTCAGAAACATTTAAATGCAGCTGCCTTCTTGCTTATAGTTTGCCTCGTTTTT    5709
GK1     GTTTGGAAAGGATTGGTTTTGTGCATATTCAGAAACATTTAAATGCAGCTGCCTTCTTGCTTATAGTTTGCCTCGTTTTT    5701
GK2     GTTTGGAAAGGATTGGTTTTGTGCATATTCAGAAACATTTAAATGCAGCTGCCTTCTTGCTTATAGTTTGCCTCGTTTTT    5709
LK1     GTTTGGAAAGGATTGGTTTTGTGCATATTCAGAAACATTTAAATGCAGCTGCATTCTTGCTTATAGTTTGCCTCGTTTTT    5693
CN1     GTTTGGAAAGGATTGGTTTTGTGCATATTCAGAAACATTTAAATGCAGCTGCCTTCTTGATTATAGTTTGCCTCGTTTTT    5732
CN2     GTTTGGAAAGGATTGGTTTTGTGCATATTCAGAAACATTTAAATGCAGCTGCCTTCTTGATTATAGTTTGCCTCGTTTTT    5740
MN1     GTTTGGAAAGGATTGGTTTTGTGCATATTCAGAAACATTTAAATGCAGCTGCCTTCTTGCTTATAGTTTGCCTCGTTTTT    5708
MN2     GTTTGGAAAGGATTGGTTTTGTGCATATTCAGAAACATTTAAATGCAGCTGCCTTCTTGCTTATAGTTTGCCTCGTTTTT    5708
B1      GTTTGGAAAGGATTGGTTTTGTGCATATTCAGAAACATTTAAATGCAGCTGCCTTCTTGCTTATAGTTTGCCTCGTTTTT    5674
B2      GTTTGGAAAGGATTGGTTTTGTGCATATTCAGAAACATTTAAATGCAGCTGCCTTCTTGCTTATAGTTTGCCTCGTTTTT    5709
SIT1    GTTTGGAAAGGATTGGTTTTGTGCATATTCAGAAACATTTAAATGCAGCTGCCTTCTTGCTTATAGTTTGCCTCGTTTTT    5658
BB1     GTTTGGAAAGGATTGGTTTTGTGCATATTCAGAAACATTTAAATGCAGCTGCCTTCTTGCTTATAGTTTGCCTCGTTTTT    5701
BB2     GTTTGGAAAGGATTGGTTTTGTGCATATTCAGAAACATTTAAATGCAGCTGCCTTCTTGCTTATAGTTTGCCTCGTTTTT    5710
IMP     GTTTGGAAACGTTTGGTTTTGTGCGGGTTCAGAAACATTTAAATGCAGCTGCCTTCTTTCTTATAGTTTACCTCGTTTTT    5710
COW     GTTTGGAAAGGATTGGTTTTGTGCATGTTCAGAAACATTTAAATGCAGCTGCCTTCTTGCTTATAGTTTGCCTCGTTTTT    5686
BUF     GTTTGGAAAGGATTGGTTTTGTGCACATTCAGAAACATTTAAATGCTGCTGCCTTCTTGCTTATAGTTTGCCTCGTTTTT    5716
NIL     GTTTGGAAAGGATTGGTTTTGTGCAGGTTCAGAAACAGTTACA------TACCTTCTTGCTTATAGTTTGCCTCATTTTT    5725
```

```
                5770       5780       5790       5800       5810       5820       5830       5840
                   .          .          .          .          .          .          .          .
GE1     T--GGTTCCTAAGTGACCCAAAAGAAAATAATACTTGGGGGAGAAAAAA-TAACATTGCTCCTCCCTCCTCTACCAATCA    5787
CE1     T--GGTTCCTAAGTGATCCAAAAGAAAATAATACTTGGGGGAGAAAAAA-TAACATTGCTCCTCCCTCCTCTACCAATCA    5779
CE2     T--GGTTCCTAAGTGATCCAAAAGAAAATAATACTTGGGGGAGAAAAAA-TAACATTGCTCCTCCCTCCTCTACCAATCA    5786
GK1     T--GGTTCCTAAGTGATCCAAAAGAAAATAATACTTGGGGGAGAAAAAA-TAACATTGCTCCTCCCTCCTCTACCAATCA    5778
GK2     T--GGTTCCTAAGTGATCCAAAAGAAAATAATACTTGGGGGAGAAAAAA-TAACATTGCTCCTCCCTCCTCTACCAATCA    5786
LK1     T--GGTTCCTAAGTGACCCAAAAGAAAATAATACTTGGGGGAAAAAAAA-TAACATTGCTCCTCCCTCCTCTACCAGTCA    5770
CN1     TTTGGTTCCTAAGTGACCCAAAAGAAAATAATACTTGGGGGAAAAAAAA-TAACATTGCTCCTCCCTCCTCTACCAATCA    5811
CN2     TTTGGTTCCTAAGTGACCCAAAAGAAAATAATACTTGGGGGAGAAAAAA-TAACATTGCTCCTCCCTCCTCTACCAATCA    5819
MN1     T--GGTTCCTAAGTGACCCAAAAGAAAATAATACTTGGGGGAGAAAAAA-TAACATTGCTCCTCCCTCCTCTACCAATCA    5785
MN2     T--GGTTCCTAAGTGACCCAAAAGAAAATAATACTTGGGGGAGAAAAAA-TAACATTGCTCCTCCCTCCTCTACCAATCA    5785
B1      T--GGTTCCTAAGTGACCCAAAAGAAAATAATACTTGGGGGAGAAAAAA-TAACATTGCTCCTCCCTCCTCTCCCAATCA    5751
B2      T--GGTTCCTAAGTGACCCAAAAGAAAATAATACTTGGGGGAGAAAAAA-TAACATTGCTCCTCCCTCCTCTCCCAATCA    5786
SIT1    T--GGTTCCTAAGTGACCCCAAAGAAAATAATACTTGGGGGAGAAAAAA-TAACATTGCTCTTCCCTCCTCTCCCAATCA    5735
BB1     T--GGTTCCTAAGTGACCCAAAAGAAAATAATACTTGGGGGAGAAAAAA-TAACATTGCTCCTCCCTCCTCTACCAATCA    5778
BB2     T--GGTTCCTAAGTGACCCAAAAGAAAATAATACTTGGGGGAGAAAAAA-TAACATTGCTCCTCCCTCCTCTACCACTCA    5787
IMP     T--GGTTCCTAAGTGACCGAAAAGAAAATAATACTTAGGGAAAAAAAGAATAACATTGCTCCCCCCTCC-----------    5777
COW     T--GGTTCCTGAGTGACCCAAAAGAAAATAATACTTCAGGGGAAAAAAA-TAACATTGTTCCTCCCTCCTCTACCAATCA    5763
BUF     T--GGTTCCTAAGTGGCCCAAAAGAAAATAATACTTGGGGGAAAAAAAA-TAACATTGCTCCTCTCTCCTCTACCATTCA    5793
NIL     T--GGTTCCTAAGTGACCCAAAAGAAAATAATACTTGGGGGGAAAAAAA-TAACATTGCTCCTCCCTCCTCTAACAATTA    5802

                5850       5860       5870       5880       5890       5900       5910       5920
                   .          .          .          .          .          .          .          .
GE1     AGATTCTAGAAGGTGACCCTGCACTTTGCATGAATGGGGGACCTAGGAACGAGAACAGGGGTGTAGCCGACTGCCTATCT    5867
CE1     AGATTCTAGAAGGTGGCCCTGCACTTTGCATGAATGGGGGACCTAGGAACGAGAACAGGGGTGTAGCCGACTGCCTATCT    5859
CE2     AGATTCTAGAAGGTGGCCCTGCACTTTGCATGAATGGGGGACCTAGGAACGAGAACAGGGGTGTAGCCGACTGCCTATCT    5866
GK1     AGATTCTAGAAGGTGGCCCTGCACTTTGCATGAATGGGGGACCTAGGAACGAGAACAGGGGTGTAGCCGACTGCCTATCT    5858
GK2     AGATTCTAGAAGGTGGCCCTGCACTTTGCATGAATGGGGGACCTAGGAACGAGAACAGGGGTGTAGCCGACTGCCTATCT    5866
LK1     AGATTCTAGAAGGTGGCCCTGCACTTTGCATGAATGGGGGACCTAGGAACGAGAACAGGGGTGTAGCCGACTGCCTATCT    5850
CN1     AGATTCTAGAAGGTGGCCCTGCACTTTGCATGAATGGGGGACCTAGGAACGAGAACAGGGGTGTAGCCGACTGCCTATCT    5891
CN2     AGATTCTAGAAGGTGGCCCTGCACTTTGCATGAATGGGGGACCTAGGAACGAGAACAGGGGTGTAGCCGACTGCCTATCT    5899
MN1     AGATTCTAGAAGGTGGCCCTGCACTTTGCATGAATGGGGGACCTAGGAACGAGAACAGGGGTGTAGCCGACTGCCTATCT    5865
MN2     AGATTCTAGAAGGTGGCCCTGCACTTTGCATGAATGGGGGACCTAGGAACGAGAACAGGGGTGTAGCCGACTGCCTATCT    5865
B1      AGATTCTAGAAGGTGGCCCTGCACTTTGCATGAATGGGGGACCTAGGAACGAGAACAGGGGTGTAGCCGACTGCCTATCT    5831
B2      AGATTCTAGAAGGTGGCCCTGCACTTTGCATGAATGGGGGACCTAGGAACGAGAACAGGGGTGTAGCCGACTGCCTATCT    5866
SIT1    AGATTCTAGAAGGTGGCCCTGCACTTTGCATGAATGGGGGACCTAGGAACGAGAACAGGGGTGTAGCCGACTGCCTATCT    5815
BB1     AGATTCTAGAAGGTGGCCCTGCACTTTGCATGAATGGGGGACCTAGGAACGAGAACAGGGGTGTAGCCGACTGCCTATCT    5858
BB2     AGATTCTACAAGGTGGCCCTGCACTTTGCATGAATGGGGGACCTAGGAACAAGAACAGGGGTGTAGCCGACTGCCTATCT    5867
IMP     ---TCCTAGACAGTGACCCTGCACTTTGCATGAATGGGGGACCTAAGAAGGAGAACAGGGGTGTAGCCGACTGCCTATCT    5854
COW     AGATTCAAGAAGGTGGCCCTGCACTTTGTATGAATGGGGGACCTAGGAACGAGAACAGGGGTGTAGCCGACTGCCTATCT    5843
BUF     AGATTCTAGAAGGTGGCCCTGCACTTGGCATGAATGGGGGACCTAGGAGCGAGAACAGGGGTGTAGCCGACTGCCTATCT    5873
NIL     GGATTCTAGAAGGTAGCCCTGCATTTTGCATGAGTGGGGGACCTAGGAACGAGAACAGGGGTTTAGCCGACCGCCTATCT    5882

                5930       5940       5950       5960       5970       5980       5990       6000
                   .          .          .          .          .          .          .          .
GE1     GTGCACGGGAGCCCATGCTACTGCTGTCATACCCGTGTGCCCGCCTTGATGTTACACAGGTTACAAGCCCTGCGACCCCC    5947
CE1     GTGCACGGGAGCCCATGCTACTGCTGTCATACCCGTGTGCCCGCCTTGATGTTACACAGGTTACAAGCCCTGCGACCCCC    5939
CE2     GTGCACGGGAGCCCATGCTACTGCTGTCATACCCGTGTGCCCGCCTTGATGTTACACAGGTTACAAGCCCTGCGACCCCC    5946
GK1     GTGCACGGGAGCCCATGCTACTGCTGTCATACCCGTGTGCCCGCCTTGATGTTACACAGGTTACAAGCCCTGCGACCCCC    5938
GK2     GTGCACGGGAGCCCATGCTACTGCTGTCATACCCGTGTGCCCGCCTTGATGTTACACAGGTTACAAGCCCTGCGACCCCC    5946
LK1     GTGCACAGGAGCCCATGCTACTGCTGTCACACCCGTGTGCCCGCCTTGATGTTACACAGGTTACAAGCCCTGCGACCCCC    5930
CN1     GTGCACGGGAGCCCATGCTACTGCTGTCACACCCATGTGCCCGCCTTGATGTTACACAGGTTACAAGCCCTGCGACCCCC    5971
CN2     GTGCACGGGAGCCCATGCTACTGCTGTCACACCCATGTGCCCGCCTTGATGTTACACAGGTTACAAGCCCTGCGACCCCC    5979
MN1     GTGCACGGGAGCCCATGCTACTGCTGTCATACCCGTGTGCCCGCCTTGATGTTACACAGGTTACAAGCCCTGCGACCCCC    5945
MN2     GTGCACGGGAGCCCATGCTACTGCTGTCATACCCGTGTGCCCGCCTTGATGTTACACAGGTTACAAGCCCTGCGACCCCC    5945
B1      GTGCACGGGAGCCCATGCTACTGCTGTCATACCCGTGTGCCCGCCTTGATATTACACAGGTTACAAGCCCTGCGACCCCC    5911
B2      GTGCACGGGAGCCCATGCTACTGCTGTCATACCCGTGTGCCCGCCTTGATATTACACAGGTTACAAGCCCTGCGACCCCC    5946
SIT1    GTGCACGGGAGCCCATGCTACTGCTGTCATACCCGTGTGCCCGCCTTGATATTACACAGGTTACAAGCCCTGCGACCCCC    5895
BB1     GTGCACGGGAGCCCATGCTACTGCTGTCATACCCGTGTGCCCGCCTTGATGTTACACAGGTTACAAGCCCTGCGACCCCC    5938
BB2     GTGCACGGGANCCCATGCTACTGCTGTCATACCCATGTGCCCGCCTTGATGTTACACAGGTTACAAGCCCTGCGACCCCC    5947
IMP     GTGCACGGGAGCCCGTGCTTCCGCTGTCACACCCATGTGCCCACCTTGATGTTACACAGGTTACAAGCCCTGCGACCCCC    5934
COW     GTGCACGGGAGCCCATGCTACTGCTGTCACACCCATGTGCCCGCCTTGATGTTACACAGGTTACAAGCCCTGCGACCCCC    5923
BUF     GTGCACGGGAGCCCATGCTACTGCTGTCACACCCACGTGCCCGCCTTGATATCACACAGGTTACAAGCCCTGCGACCCCC    5953
NIL     GTGCACGGGAGCCCATGCTACCGCTGTCACACCCATGTGCCCACCTTGATGTTACACAGGTTACAAGCCCTGCGACCCCC    5962
```

GE1	AGGTAATC	5955
CE1	AGGTAATC	5947
CE2	AGGTAATC	5954
GK1	AGGTAATC	5946
GK2	AGGTAATC	5954
LK1	AGGTAATC	5938
CN1	AGGTAATC	5979
CN2	AGGTAATC	5987
MN1	AGGTAATC	5953
MN2	AGGTAATC	5953
B1	AGGTAATC	5919
B2	AGGTAATC	5954
SIT1	AGGTAATC	5903
BB1	AGGTAATC	5946
BB2	AGGTAATC	5955
IMP	AGGTAATC	5942
COW	AGGTAATC	5931
BUF	AGGTAATC	5961
NIL	AGGTAATC	5970